# This Page Is Inserted by IFW Operations and is not a part of the Official Record

### **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

### IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

### (19) World Intellectual Property Organization International Bureau





#### (43) International Publication Date 9 August 2001 (09.08.2001)

#### PCT

### (10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

- - -

(21) International Application Number: PCT/US01/03800

(22) International Filing Date: 5 February 2001 (05.02.2001)

(25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 .3 February 2000 (03.02.2000) US 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US - 09/496,914 (CIP)
Filed on 3 February 2000 (03.02.2000)
US 09/560,875 (CIP)
Filed on 27 April 2000 (27.04.2000)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

01/57188 A

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

#### NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

#### 1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

#### 2. BACKGROUND

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

#### 3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, \* corresponds to the stop codon.

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

5

10

15

20

725

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

10

15

20

25

30

35

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

15

20

25

30

35

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

#### 4. DETAILED DESCRIPTION OF THE INVENTION

#### **4.1 DEFINITIONS**

10

15-

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

5

10

20

25

30

35

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4<sup>20</sup> possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match  $(1 \div 4^{25})$  times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that. have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (*i.e.*, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (*i.e.*, washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

15

5

10

#### 4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

5

10

- 15

20

25

30

35

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

#### 4.3 ANTISENSE-

5

10

15

20

25

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inque et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inque et al. (1987) FEBS Lett 215: 327-330).

#### 4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

5

10

. 15

20

25

30

35

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

5

10

15

20

#### 4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., Basic Methods in Molecular Biology (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells-under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

20

25

30

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

5

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane-bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

10

15

20

25

30

35

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat<sup>TM</sup> kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl<sup>TM</sup> or Cibacrom blue 3GA Sepharose<sup>TM</sup>; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

- 15

20

25

30

35

5

10

## 4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

#### 4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

10

15

20

25

30

35

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (*i.e.*, glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

#### 4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered in vivo to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.9 TRANSGENIC ANIMALS

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

5

10

15

20

25

30

35

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

### 4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

# 4.10.1 RESEARCH USES AND UTILITIES

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

#### 4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

# 4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

35

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

# 4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

#### 4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) . as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

30

10

### 4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

# 4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

5

10

15

20

25

30

35

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

10

15

20

25

30

35

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

5

10

15

20

25

30

35

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and  $\beta_2$  microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

5

10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

# 4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

# 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

# 4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

# 4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

5

10

15

20

25

30

35

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

### 4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

5

10

15

20

25

30

35

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

# 4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

<del>2</del>5

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

# 4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

35

30

5

10

15

20

25

# 4.10.15 ANTI-INFLAMMATORY ACTIVITY

5

10

15

20

35

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

## 4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a
therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see
Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

# 4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

5

10

15

20

25

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;
- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular
   neurotoxins; and
  - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

30

35

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
  - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

#### 4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

#### 4.10.19 **IDENTIFICATION OF POLYMORPHISMS**,

10

15

25

30

35

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all-generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

#### 4.10.20 ARTHRITIS AND INFLAMMATION

5

10

15

20

25

30

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

## 4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

#### 4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

# 4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

#### 4.12.1 ROUTES OF ADMINISTRATION

5

10

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

#### 4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

5

10

15

20

25

30

35

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

35

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

5

10

20

25

30

35

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

5

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and 10 tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl 15 cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

20

25

30

35

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

10

15

#### 4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC<sub>50</sub> as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD<sub>50</sub> and ED<sub>50</sub>. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01  $\mu$ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1  $\mu$ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

#### 4.12.4 PACKAGING

5

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

## 4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain,  $F_{ab}$ ,  $F_{ab}$  and  $F_{(ab)2}$  fragments, and an  $F_{ab}$  expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG<sub>1</sub>, IgG<sub>2</sub>, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

## 5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corvnebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

## 5.13.2 Monoclonal Antibodies

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol.</u>, <u>133</u>:3001 (1984); Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

## 5.13.2 Humanized Antibodies

10

15

20

25

30

35

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

#### 5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, <u>J. Mol. Biol.</u>, <u>227</u>:381 (1991); Marks et al., <u>J. Mol. Biol.</u>, <u>222</u>:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (<u>Bio/Technology 10</u>, 779-783 (1992)); Lonberg et al. (<u>Nature 368</u> 856-859 (1994)); Morrison (<u>Nature 368</u>, 812-13 (1994)); Fishwild et al., (<u>Nature Biotechnology 14</u>, 845-51 (1996)); Neuberger (<u>Nature Biotechnology 14</u>, 826 (1996)); and Lonberg and Huszar (<u>Intern. Rev. Immunol. 13</u> 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse<sup>TM</sup> as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

## 5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of Fab expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal Fab fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an Fab fragment generated by reducing the disulfide bridges of an Fab fragment; (iii) an Fab fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) Fab fragments.

## 5.13.5 Bispecific Antibodies

10

15

20

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

10

15

20

25

30

35

(

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

## 5.13.6 Heteroconjugate Antibodies

5

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

## 5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

#### 5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include  $^{212}$ Bi,  $^{131}$ I,  $^{131}$ In,  $^{90}$ Y, and  $^{186}$ Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

## 4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

### 4.15 TRIPLE HELIX FORMATION

5

10

20

25

30

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem.

56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

#### 4.16 DIAGNOSTIC ASSAYS AND KITS

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

#### 4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

#### 4.18 SCREENING ASSAYS

5

10

- 15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
  - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

10

15

20

25

30

35

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA-binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

## 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

10

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

#### 4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm<sub>7</sub>), is then added to a final concentration of 10 mM 1-MeIm<sub>7</sub>. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm<sub>7</sub>, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

# 4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

5

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviII, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviJI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (CviJI\*\*), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald et al. (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a CviJI\*\* digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that CviJI\*\* restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

# 4.22 PREPARATION OF DNA ARRAYS

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm<sup>2</sup> and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

## 5.0 EXAMPLES

5

10

15

20

25

30

#### 5.1 EXAMPLE 1

# Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

## 5.2 EXAMPLE 2

### Novel Contigs

5

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from <a href="http://fasta.bioch.virginia.edu">http://fasta.bioch.virginia.edu</a>) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

5

```
#####
                                             ##
####
          # #
# #
####
# #
# #
# #
                                              #
# ###
## ;
                 #
                                                           ### ###
                                                            ##
                                                                  #
                       ####
   ###
                                              ##
                             ##
                                   ####
                                                                  #
                                         ##
                                              ##
                           ##
   #
                        ## ## #####
         #####
                                             ## ###
```

Job : 110

Date: 5/19/2004 Time: 11:24:54 PM

```
##
  ####
         #####
                                      #
# ###
    ##
          #
                             #####
# #
                                                ### ###
               #
                        ##
                                                 #
                                      ##
                                                      #
          ####
# #
                         ##
    ###
                                      ##
             #
                              ####
                                                 #
                                                      #
###
                                                  # #
                    ##
                         #
          #
                              #
                                   #
                                      ##
                                                  # #
                                            #
               #
                        ##
         #####
                     ## ## #####
 ###
                                     ## ###
                                                 ###
```

Job: 120

Date: 5/19/2004 Time: 11:57:58 PM

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
			976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
			217 225 238 271 317 404 446 469 503
	}	}	513-514 535 550 564 573 666-669 798
	Ļ		898 910 927 976 1067 1083 1085 1178
	1		1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
	·	ì	147 188 197 208 225 227-239 250 300-
	1	1	303 312 316 328-331 340 357-362 374
	1	· ·	380 384-391 408 414 446 448 464-467
	1	ļ	483 488 495-496 505 512 521 535 550
	1	}	566 571 577 585 590 594 598 634 641
	<b>,</b>	}	658 666 683 725 742 764 767 786 801
	}		805 810 823 826 829 831 836 841 887-
		1	923 927 934 943 950-951 963 976 995
	1		1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
			1128 1142 1162 1181-1192 1199 1204
-			1218-1219 1225 1232 1253 1267 1271-
			1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
		. "	566 596 663 670 746 798 816-819 876
	j		892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes		1	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
lunional grand	0.00		240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
		†	1003 1067-1070 1118 1156 1193-1200
		}	1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
			118 129 132 138 151 158-163 182 195-
		(	203 215 217 238 264 269 353 384 398
		{	408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
			874 891 898 919 926-927 976 988
Ì			1021 1037 1041 1062 1067 1071 1080
			1083 1093 1122 1131 1185 1201 1254
			1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
adult kidney	dibeo		107 111 119-120 138 151 157 215 217-
	1		
	1		218 238 250 264 294 304 384 404 440
			446 454 477 504-505 509 514 518-519
			446 454 477 504-505 509 514 518-519 535 537 564 574-583 620-627 639 653 673-675 705 753 789 831 844 851 859
			446 454 477 504-505 509 514 518-519 535 537 564 574-583 620-627 639 653 673-675 705 753 789 831 844 851 859 877 909 918 927 956 963 976 1067
A.06 V. J.	Invita	AKT002	446 454 477 504-505 509 514 518-519 535 537 564 574-583 620-627 639 653 673-675 705 753 789 831 844 851 859 877 909 918 927 956 963 976 1067 1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	446 454 477 504-505 509 514 518-519 535 537 564 574-583 620-627 639 653 673-675 705 753 789 831 844 851 859 877 909 918 927 956 963 976 1067 1074 1083 1095 1178 1302 1331 1335 11-12 41 49 111-112 215-217 294 316
adult kidney	Invitrogen	AKT002	446 454 477 504-505 509 514 518-519 535 537 564 574-583 620-627 639 653 673-675 705 753 789 831 844 851 859 877 909 918 927 956 963 976 1067 1074 1083 1095 1178 1302 1331 1335 11-12 41 49 111-112 215-217 294 316 446 487 564 575 844 868 910 927 976
adult kidney	Invitrogen	AKT002 ALG001	446 454 477 504-505 509 514 518-519 535 537 564 574-583 620-627 639 653 673-675 705 753 789 831 844 851 859 877 909 918 927 956 963 976 1067 1074 1083 1095 1178 1302 1331 1335 11-12 41 49 111-112 215-217 294 316

lissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			518 537 545 549 580 582 592 594 634
	<b>\</b>		640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
ymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
<b>,</b> 1			545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
, , , , , , , , , , , , , , , , , , , ,			519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
	1	1	1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
uctuit iivoi			505 657 675 714 753 832 844 941-942
	1	į.	976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult nver	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
adult ovary	mvidogen	7.0 7 00 7	104 111 120 122-125 138 140 143-149
			151 188-190 207-212 215-217 238 264
			316 384 409 440 445-446 496 504 512
		1	514 518-519 535 537 549-550 564 566
		1	571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771
			815 833 842-865 868 875-876 918 926-
			927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
		ļ	1124 1131 1144 1174 1224 1268 1331
			1335
		APL001	102 217 238 537 641 700
adult placenta	Clontech		663 851 1048
placenta	Invitrogen	APL002	8 45 74 111 132 140 151 185 217 238
adult spleen	GIBCO	ASP001	294 414 446 477 504 514 534 545 549
	ļ	)	592 722 873 883 952 976 1041-1042
		<b>\</b>	
			1083 1093-1094 1152 1224 72 107 111 113 126 140 151 183 215
testis	GIBCO	ATS001	238 446 497 537 642 701-706 811 877
			927 962 976 1083 1117 1131 41 151 191 402-405 409 414 496 545
adult bladder	Invitrogen	BLD001	592 607 706 873 952 1178 1329-1335
			392 607 700 873 932 1176 1329-1333
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
		l	147 151 164-174 213-215 238 305-307
		1	374 404 446 460 466 516 519 534 538-
	.)		541 544-546 549-554 566 584 586 592
		ì	596 607 610 628-629 643-645 652 707-
ļ	}	}	708 774-789 844 866-871 873 919 927
			952 963 976 998 1034 1042 1064 1083
		ì	1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
			210 317 510-511 545 549 581 598 628
		1	638 724 766 789 844 860 868 873 919
		•	
		·	927 952 963 968 976 1042 1111 1141
		•	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083
	Clontech Invitrogen	BMD004 CLN001	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592
bone marrow adult colon			927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268
			927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337
adult colon	Invitrogen	CLN001	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332-
			927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332-
adult colon	Invitrogen	CLN001	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337
adult colon	Invitrogen	CLN001	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332-335 365-367 392-401 440 466 470-471 518 537 597 629 832 877 927 976 1006
adult colon	Invitrogen	CLN001	927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332-335 365-367 392-401 440 466 470-471

Tiama Origin	RNA Source	Hyseq Library Name	SEO ID NOS:
Tissue Origin endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
епиотиены сень	Sualegene	ISD 1001	138 151 204-206 215-217 238 269 316
			414 433 505 510 513 550 555 580 582
	{		596 675 722 745 798 814 836-841 851
			918 976 1041 1043 1073 1083 1131
ĺ			1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic	223.200	
of chromosome 8	Research		<u> </u>
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		1
of chromosome 8	Research		1
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic		
of chromosome 8	Research	}	
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic		1
of chromosome 8	Research		)
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
Town order	0.0		225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
		,	829 836 859 909 927 943 947 963 1057
			1067-1068 1104 1135-1140 1162 1206-
			1207 1235 1268 1288 1307-1308 1319
	Ì		1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
			535 683 761 798 820-827 844 876 909
	ļ	ĺ	963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
			550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
			859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
1	University	,	69-71 74 77 79 87-90 101 107 110-111
			114 120 128-131 138 140 147 150-155
			197 210 215 217 225 238 312 367 384
			414 440 446 460 468 483 496 504-507
			511-515 518-519 523 533-535 537 541
	1		544-545 547-550 555-560 564 566 571
			577 582 585-586 598 636 646-647 649
	ĺ		652 664 698 709-710 714 722-723 731
<u> </u>	1		735-736 746-753 761 784 798 823 829
			832 844 851 858-859 868 873 876 898
	}		927 943 949 952 963 976 984 1002
		1	1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174
			1093 1116 1120 1129 1131 1144 1174
		1	
	<del></del>	Tr. coop	1319 8 36-37 41-46 49 54 64 71 74 79 101
fetal liver-spleen	Columbia	FLS002	111 120 129 147 207 210 215-216 238
	University		250 330 353 359 366 383-384 414 478
			505 508-509 511 515-524 534-535 537
		}	544-545 564 566 571 577 591 598 638
L	_ \		סכט 270 ברני דוב בוב דטב ברנידרב ן

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
<u> </u>			663 671 698 714 722 725 727 751 798
	{		851 859 873 876 909 927 949 952 983-
			984 1002 1023 1042-1044 1085 1095
	1		1131 1144 1178 1199 1233 1240-1270
			1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
20000 227 02			580 722 730 749 844 918 943 976 1051
			1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
			425 535 537 577 598 614 836 857 1141
	<u> </u>		1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
			225 264 316 405 422-429 488-494 496
			519 534-535 537 566 675 732 859 876- 877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
•	1	1	1281
<u> </u>		FSK002	537 812
fetal skin	Invitrogen BioChain	FSP001	87 549
fetal spleen umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
umbilical cord	- BioChain	rocooi	316 446 495-503 519 521 534-535 537
•			582 634 691 877 883 927 944-950 963
			976 1001 1075 1142-1143 1171 1218
			1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
iciai oram	GIECO	122001	138 145 151 188 197 207 215 238 264
			271 294 316 367 414 440 446 466 504
			513-514 535 542-543 550 564 571 596
		ļ	635 648-654 675 711-715 722-723 798
			832 872 876 883 927 976 1095 1144
			1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
	1		299 308-310 371-373 462 476 504 511-
			513 533 537 564 566 571 655-657 662 683 716-720 723 752 790-803 829 832
9			858-859 876 898 909 949 976 1045-
+		· ·	1047 1076-1087 1090 1093 1116 1122
		}	1144 1209-1213 1225 1233 1256 1319
		ì	1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
miant orani	University	152003	519 566 655 714 794 918 943 976 1067
	Cimversity	}	1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
	University		
infant brain	Columbia	IBS001	51 111 376 474 790 876 949 1144 1204
	University		1221
lung, fibroblast	Strategene	LFB001	151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156
			215 217 269 280 296 337 374-375 384
	1		404 446 454 475-480 498 514 518-519
	Ì		522 537 545 564 577 597 653 658 705
			721-724 754-756 779 859 868 872-874
i .	í	1	876-877 919 927 949 951-952 959 976
	i	1	
			1002 1042 1048-1053 1076 1083 1088- 1089 1131 1144-1147 1216-1218 1229

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550 634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
leukocyte	GIBCO	Locooi	147 151 212 215 218 238 252 288 312-
i			314.316.338.359.408.427.443-447.505
		i	510 512 514 518 534 545 549-550 561
		ł	564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
		}	836 841 859 866 873-874 882-883 918-
		,	919 927 943 952 963 976 1042 1076
		1	1083 1090 1148 1152 1168 1195 1219-
	ļ		1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
		1	657 660 729 873 883 927 952 963 1008
	ĺ		1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL	Cioniccai		919 929 939 952 976 1071 1118 1218
1424	}	1	1235 1245
	Tourismonom	MMG001	8-10 40-41 49 73 80 114 138-140 147
mammary gland	Invitrogen	INTIATOROIT	217 250-256 264 297-299 305 377-378
			398 446 481-486 505 512 537 545 549
		•	571 592 725 730-733 816 829 836 844
		ĺ	868 873 876-877 898 926 943 951-960
	:		963 976 995 1034 1042 1048 1054-
	}		1055 1076 1083 1091 1093 1116-1117
	ļ		1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
•		·	1319
retinoid acid induced	Strategene	NTR001	74 225 976
neuronal cells			
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
		·	1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
			1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
, ,			976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
			545 592 660 789 836 866 873 927 952
			963 967-978 1042 1120 1152 1223-
, ·		•	1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
spinal cord	Cioniecii	31 C001	270 343-344 353 379 516 537 566 740
	1		828 927 976 979-994 1092 1153-1159
	ļ	1	
		CDY 01	1225 1250 698 859 1042
adult spleen	Clontech	SPLc01	210 238 271-272 537 580 705 918 952
stomach	Clontech	STO001	995 1171
		(TTT 4 000	61 219-220 273-276 312 315 330 596
thalamus	Clontech	THA002	963 996-1007 1059 1093 1160-1162
		CTT 70 500 ;	8 120 151 208 221 316-317 353 639
thymus	Clonetech	THM001	
		}	750 867 874 878-881 927 963 1023
			1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306
	1		317-319 336 340 359 380 398 446 448
	1	1	463 512 519 545 554 587 598 698 724
1		ŀ	725 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
Thous origin		1	976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
			1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
	1	ļ	320-326 345-352 361 381-382 446 467
	1		483 514 534 549-550 564 578 602 649
	<b>\</b>		844 882-883 927 950 956 976 1008-
	}		1028 1076 1083 1117-1120 1142 1163-
	}		1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
			545 592 611 873 883-884 927
{			952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
,			885-886 976 1001 1032-1033
			1232

## TABLE 2

SEQ	Accession	Species	Description	Smith-	%
D C	No.			Waterman	Identity
NO:	İ			Score	1.00
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	1,29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
<del>5</del>	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Hurnan secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ D NO:	Accession No.		Description	Smith- Waterman Score	% Identity
9	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
<del></del>	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
1	G02372	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	58	32
32		Homo sapiens	Membrane-bound protein PRO1152.	2457	98
33	Y66688		Human secreted protein sequence SEQ ID	348	95
34	Y87071	Homo sapiens	NO:110.	182	48
35	U15131	Homo sapiens	p126	982	90
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.		
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
40	AF132969	Homo sapiens	CGI-35 protein	228	68
	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
42			mini-collagen	105	35
43	X61048	Hydra sp.	hydroxyproline-rich protein	110	31
44	M76546	Helianthus annuus		139	70
45	U82288 	Caenorhabditi s elegans	Rac-like GTPase		
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus musculus	SPR2B protein	72	56
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
	1460041	Homo sapiens	protein-tyrosine phophatase	165	41
56	M68941		c390E6.1 (chloride channel 7)	338	76
57 58	AL031600 AF011417	Homo sapiens Mus	putative pheromone receptor	143	55
59	AF 167320	musculus Mus	zinc finger protein ZFP113	558	68
	110000	musculus	interferon regultory factor 7	263	96
61	X07984	Mus	protein-tyrosine kinase	297	69
-70	1/00001	musculus	Human secreted protein clone cb98_4.	791	98
62	Y29861	Homo sapiens		485	65
63 64	U35376 AF265555	Homo sapiens Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	785	74
		<del></del>	APOLLON SEO ID NO. 7064	88	95
65	G03883	Homo sapiens			54
66	AF177390	Manduca sexta	antennal specific membrane protein AMP	274	
67	AB040800	Homo sapiens		614	100
68	AF030027	Equine herpesvirus 4	24	213	26
69	G02965	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70	W75770	Homo sapiens		1144	98
71	AB011135	Homo sapiens		239	76
72	AB014885	Halocynthia	HrPOPK-1	813	78
73	AF045454	Cavia porcellus	phospholipase B	955	73
		i matteniik	1	1	

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO:		musculus			<del> </del>
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
3	100020	norvegicus			
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	351	54
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha- 1-I (hCavT3).	1357	99
79	Y14591	Human papillomaviru s type 68	APM-1 protein	767	100
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis thaliana	protein arginine N-methyltransferase-like protein	359	65
82	L46815	Mus musculus	DNA binding protein Rc	895	75
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	538	71
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272 7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid seguence SEO ID NO:100.	156	48
88	AJ225124	Mus musculus	hyperpolarization-activated cation channel, HAC3	487	95
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK10	401	53
94	X13292	Trypanosoma brucei	GPI-phospholipase C (AA 1 - 358)	151	37
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	675	48
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence difference at residue 58	160	60
102	U22829	Mus musculus	P2Y purinoceptor	264	42
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled receptor-B3.	516	99
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119.	343	57
106	AF169312	Homo sapiens		212	67
107	AF116657	Homo sapiens		74	52
108	AE000401	Escherichia coli	sialic acid transporter	587	96
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	182	94
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	464	85
112	Y94939	Homo sapiens		274	51
113	AF016365	Homo sapiens		301	71
114	AC007956	Homo sapiens		520	75
115	M83738	Homo sapiens	protein-tyrosine phosphatase	251	92
116	AL157952	Homo sapiens	dJ875K15.1.1 (ets homologous factor (ets- domain transcription factor ESE-3A, isoform 1))	484	91
117	W18084	Homo sapiens		546	87

EQ	Accession	Species	Description	Smith- Waterman	% Identity
D	No.	, <u>,</u>		Score	1
<u>10:</u>				407	62
18	L41816	Homo sapiens	cam kinase I	627	93
19	AJ006710	Rattus norvegicus	phosphatidylinositol 3-kinase		
20	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1646	94
21	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
22	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
23	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
24	U88167	Caenorhabditi	contains similarity to C2 domains	219	29
		s elegans	·		
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit 4	693	90
126	AB021861	Mus musculus	apoptosis signal-regulating kinase 2	153	65
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IFI16b	496	67
131	AF201734	Mus musculus	testis specific serine kinase-3	800	87
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73	1157	87
			clone HSQEL25.		
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674_2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95% similarity to P49205 (PID:g1345860)	5041	99
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis thaliana	DNA-damage inducible protein DDI1-like	147	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia coli	HrsA	818	90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens		940	99
152	R95332	Homo sapiens		392	61
153	AF151859	Homo sapiens		370	92
154	X66957	Homo sapiens		489	81
155	Y16355	Homo sapiens		432	92
156	G00857	Homo sapiens		349	78
157	AF159455	Mus	zinc finger protein	352	74
1.55		musculus	interleukin-1 receptor-associated kinase	537	76
158 159	L76191 AP001743	Homo sapiens	putative gene, ankirin like, possible dual	670	98
160	AJ250425	Rattus norvegicus	specifity Ser/Thr/Tyr kinase domain  Collybistin I	556	74
1					

EQ D	Accession No.	Species	Description	Smith- Waterman Score	% Identity
10:	Z22968	Homo sapiens	M130 antigen	610	100
62	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
63		Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
64 65	AF055636 AF160798	Rattus norvegicus	calcium transporter CaT1	700	96
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
68	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinae gen. sp.	diacylglycerol kinase eta	481	82
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Home sapiens	Human secreted protein encoded by gene No. 123.	653	99
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196_4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in codon)	710	99
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	90
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	49
195	W95631	Homo sapiens	hi968 2.	382 680	99
196	AF255614	Rattus norvegicus	scaffolding protein SLIPR	300	41
197	AC021640	Arabidopsis thaliana	putative phosphatidate phosphohydrolase	316	43
198	AF073967	Mus musculus domesticus	olfactory receptor		
199	W01730	Homo sapiens		617	98
200	AF117948	Homo sapiens		625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79
205	S81752	Homo sapiens	DPH2L=candidate tumor suppressor gene	375	100

SEQ D	Accession No.	Species	Description	Smith- Waterman	% Identity
10:				Score	
			{ovarian cancer critical region of deletion}		
06	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
			protein, calphotin.	1240	99
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	
212	U81035	Rattus norvegicus	ankyrin binding cell adhesion molecule	471	69
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	933	93
		musculus			<del> </del>
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein	563	78
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
	X75756	Homo sapiens	protein kinase C mu	314	81
218		Homo sapiens	Membrane-bound protein PRO1100.	770	98
219	Y66723			567	40
220	D88577	Mus musculus	Kupffer cell receptor		
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding kinase	636	96
223	AL136527	norvegicus Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
			11)	690	99
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	703	68
225	AF030430	Mus musculus	semaphorin VIa	/03	08
226	AE000218	Escherichia coli	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus	GTP-binding like protein 2	265	88
229	AF122924	Musculus Xenopus	Wnt inhibitory factor-1	316	40
		laevis	5 CEO ID NO. 7086	229	100
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	265	92
231	X98260	Homo sapiens	M-phase phosphoprotein 11		
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific phospholipase-D.	290	100
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
236			Embryo Brain Kinase	460	62
237	X81466	Mus musculus			
238	U64857	Caenorhabditi s elegans	similar to the BPTI/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor (TFPI)	284	33
239	AJ250840	Mus musculus	serine/threonine protein kinase	739	63
240	AJ223472	Mus musculus	transcription elongation factor TFIIS.h	222	38
241	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein sequence SEQ ID NO:18.	353	52
345	A 75 / C0204	Unma assisse		591	99
242	AF169301	Homo sapiens		667	93
243	L22022	Rattus norvegicus	orphan transporter v7-3		
244	AF016191	Rattus norvegicus	potassium channel	1043	98
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		Lane adjum calaium natarajum ayahangan	645	98
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	497	98
246	Y29868	Homo sapiens			
247	AF180475	Homo sapiens		188	83
248	Y17227	Homo sapiens	Human secreted protein (clone yal-1).	690	99
249	AF250910	Manduca	death-associated small cytoplasmic leucine-rich	182	31

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
		sexta	protein SCLP		<u> </u>
250	AF192756	Kaposi's sarcoma- associated herpesvirus	Orf73	134	34
251	AB022694	Homo sapiens	MOK protein kinase	209	83
52	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus musculus	DNA binding protein Rc	251	67
54	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus musculus	Citron-K kinase	1201	98
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus cuniculus	Phospholipase	368	80
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus musculus	L-periaxin	430	72
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus norvegicus	SLIT-2	198	40
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor (GPCR).	636	99
266	U27269	Mus musculus	sodium glucose cotransporter	204	56
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus norvegicus	putative taste receptor TR1	209	39
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus pyogenes	Fc-gamma receptor	129	26
271	AB009883	Nicotiana tabacum	KED	109	26
272	AF137367	Mus musculus	VPS10 domain receptor protein SORCS	899	97
273	L34938	Rattus norvegicus	ionotropic glutamate receptor	460	86
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)	188	74
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	173	94
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278 279	AB046851 AC008075	Homo sapiens Arabidopsis	KIAA1631 protein Contains PF 00069 Eukaryotic protein kinase	157	96
		thaliana	domain.	1	<del> </del>
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICEI	605	76
283	AF156530	Mus musculus	ETS-domain transcriptional repressor PE1	1	
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate reading frame protein.	647	100
285	Y73402	Homo sapiens	sequence SEQ ID NO:26.	300	90
286	AF016411	Homo sapiens	KCNA3.1B	137	100
287	W89253	Homo sapiens	Human ALP.	688	97
288	AF112886	Bos taurus	differentiation enhancing factor 1	750	96
289	AF113131	Homo sapiens		367	100
290	U52111	Homo sapiens		698	100
291	AF026504	Rattus	SPA-1 like protein p1294	603	89

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
		norvegicus			
92	AF102854	Rattus	membrane-associated guanylate kinase-	124	53
		norvegicus	interacting protein 2 Maguin-2		
293	X99211	Drosophila melanogaster	ubiquitin-specific protease	143	38
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	185	94
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577_1.	568	84.
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	182	97
299	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	605	69
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases; Method: conceptual translation supplied by author	428	72
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus musculus	membrane glycoprotein	199	41
308	AF255614	Rattus norvegicus	scaffolding protein SLIPR	639	88
309	S79463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium discoideum	calcium binding protein	151	36
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	744	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:g1736918)	197	38
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and GENEWISE)	278	38
316	U70209	Mus musculus	polycystic kidney disease I protein	165	38
317	AF109643	Rattus norvegicus	coxsackie-adenovirus-receptor homolog	223	38
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma vivax	activated protein kinase C receptor homolog	141	38
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate	232	97
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	306	88
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37)	214	97
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens		540	78
328	G02292	Homo sapiens		721	99
329	AF168990	Homo sapiens		877	99
330	S67984	Homo sapiens		581	80
331	X13916	Homo sapiens		2823	98
332	Y87330	Homo sapiens		1127	100
333	Y28503	Homo sapiens		320	98
334	AC002563	Homo sapiens	0.704	327	93

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO:		<del></del>	similarity to P49205 (PID:g1345860)		
	3707247	Homo sapiens	Human signal peptide containing protein HSPP-	1111	67
35 	Y87347		124 SEO ID NO:124.	193	75
36	AF006466	Mus musculus	lymphocyte specific formin related protein		
37	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
38	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
39	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
40	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
41	AE000497	Escherichia	L-idonate transcriptional regulator	928	98
342	D90855	coli Escherichia	glycerol-3-phosphate dehydrogenase (EC	769	99
343	D85613	coli Escherichia	1.1.99.5) chain A, anaerobic membrane component	399	100
344	M93239	coli Escherichia	transmembrane protein	232	100
		coli		759	99
345	M60177	Escherichia coli	enterobactin		
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia	49 kd protein	1193	96
349	L10328	Escherichia	similar to drug resistance translocases	340	90
350	X69942	coli   Mus	enhancer-trap-locus-1	560	82
351	AF239613	musculus Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
352	D90777	Escherichia	activated potassium channel 3-hydroxybutyry!-CoA dehydrogenase (EC	577	100
352	D90777	coli	1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).		
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
	AF119226	Homo sapiens		1788	100
357 358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID	165	100
250	100122	Homo sapiens	NO:258. beta-fibrinogen	233	93
359	J00132			128	70
360	G03789	Homo sapiens Homo sapiens		108	40
361 362	R28916 U16655	Rattus	phospholipase C delta-4	649	65
262	C02110	norvegicus Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
363	G03119	Gallus gallus	chicken brain factor-2	104	34
364	U47276			183	65
365	G03789	Homo sapiens		118	46
366	G04091	Homo sapiens		564	75
367	X98258	Homo sapiens		3387	99
368	AL021366	Homo sapiens		92	59
369	U70932	Peromyscus leucopus	reverse transcriptase		
370	X86400	Homo sapiens	like	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	(1.40.4506)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
	ı	Homo sapiens		172	55

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:				Score	
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus musculus	GTP binding protein	1456	91
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
380			LAK-4p	530	43
381	AB002405	Homo sapiens		115	44
382	U64830	Dictyostelium discoideum	protein tyrosine kinase		<u> </u>
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
		Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
388	G04072			197	51
389	M12140	Homo sapiens	envelope protein	461	177
390	AJ293309	Homo sapiens	NHP2 protein		94
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228 6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	171	34
206	002020	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
396 397	G03930 AB032904	Hylobates	dopamine receptor D4	105	35
		syndactylus	2 (074 (2)	861	.85
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)		92
399	Ÿ91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	
400	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number Z21513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus musculus	ADP-ribosylation factor-directed GTPase activating protein isoform b	545	89
405	X92887	Human endogenous retrovirus K	pol/env	162	30
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
407	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone HTSEV09.	2004	99
411	AB043953	Mus musculus	Chat-H	2628	82
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
413	U10542	Pan troglodytes	MHC class I A	265	71
414	A E1 55007	Homo sapiens	NY-REN-7 antigen	850	95
414	AF155097			88	48
415	G03203	Homo sapiens	1 7700 000 1 0 5	266	89
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.		
417	W27651	Homo sapiens		481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens		1446	96
422	AC008075	Arabidopsis thaliana	F24J5.4	112	35

SEQ D	Accession No.	Species	Description	Smith- Waterman Score	% Identity
<b>VO:</b>				1090	100
23	AF231705		Alu co-repressor I	6268	97
24	AF234887	Homo sapiens	FLAMINGO 1	1961	99
25	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 191.		
26	AB009288	Homo sapiens	N-copine	635	98
27	L12392	Homo sapiens	Huntington's Disease protein	16080	99
28	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
29	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
30	Y84441	Homo sapiens	Amino acid sequence of a human RNA- associated protein.	2074	100
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
132		Lycopersicon	extensin-like protein	613	48
433	AF159296	esculentum	Human breast cancer related protein BCRB2.	135	44
434	W48351	Homo sapiens		3442	97
435	X73874	Homo sapiens	phosphorylase kinase	268	74
436	AF161426	Homo sapiens	HSPC308	1055	52
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.		
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
	X98330	Homo sapiens	ryanodine receptor 2	9356	99 -
445		Homo sapiens	PRO2738	227	49
446	AF116712		sphingosine kinase type 2 isoform	576	99
447	AF245447	Homo sapiens	membrane-type serine protease l	2630	94
448	AF133086	Homo sapiens		817	93
449	U87305	Rattus norvegicus	transmembrane receptor UNC5H1		99
450	AF081249	Homo sapiens	JAW1-related protein MRVI1A long isoform	4568	
451	AC005498	Homo sapiens	R31665_1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1 (CIRP-1).	192	67
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by gene 62.	106	40
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
457	M19155	Plasmodium	S-antigen precursor	110	36
470	[411,112]	falciparum		Į	
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
		Homo sapiens	to the second of the same AA	149	43
460	Y02693	Homo sapiens	clone HTDAD22.	184	54
461	¥14482	Homo sapiens	gene 17.	135	47
462	Y53005		sequence SEQ ID NO:16.	109	33
463	X84960	Triticum aestivum		1781	85
464	W19919	Homo sapiens		502	59
465	AF189764	Mus musculus	alpha/beta hydrolase-1	J	
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens		1172	99
1		Home conien	2000	149	52
468	G02872	Homo sapiens		5832	97
469	AJ000008 X70922	Homo sapiens Mus	neurotoxin homologue	118	47
		musculus	Human secreted protein, SEQ ID NO: 7878.	198	75
471	G03797	Homo sapiens			57
472	Y36705	Homo sapien:	e i Fragment of numan secreted protein encoded by	12	1 31

SEQ	Accession	Species	Description	Smith- Waterman	% Identity
D .	No.				Incitnità
40:	<u> </u>	<u> </u>		Score	<del> </del>
		L	gene 62.	220	100
73	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	97
174	Y07007	Homo sapiens	Breast cancer associated antigen precursor sequence.	1013	9/
175	W93254	Homo sapiens	Human ESRPI protein.	943	80
176	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	202	60
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	3427	92
4/2	AL IVE	musculus	i 1 1 D Imgo. vomania prosperanto inicia		
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
461.			SYTAXI.		
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483	AF210651	Homo sapiens	NAG18	124	59
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein 3	149	73
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-l	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	197	67
495	AC005175	Homo sapiens	R31449 3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
437	12030237	familiaris		1	
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus	reverse transcriptase	213	52
477	0,0333	maniculatus	100000000000000000000000000000000000000		Ì
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124_3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135_9.	986	70
507		Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
	Y86265		NO:180.		
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light polypeptide kinase))	184	92
509	U43360	Peromyscus maniculatus	reverse transcriptase	97	62
510	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
511	W79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
512	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
513	AJ133439	Homo sapiens	GRIP1 protein	2151	100
514	AE003456	Drosophila	CG6393 gene product	259	42 .
515	Z17206	melanogaster Xenopus	p46XlEg22	128	40
		laevis		1764	
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
517	G03797	Homo sapiens		92	40
	AF151083	Homo sapiens		444	98
518 ·					
518 · 519 520	S80864 X92485	Homo sapiens Plasmodium	cytochrome c-like polypeptide	170	61

SEQ	Accession	Species	Description	Smith-	%
D	No.	·		Waterman	Identity
TO:				Score	<u> </u>
21	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
22	AF121857	Homo sapiens	sorting nexin 7	259	40
23	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
24	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	253	73
25	AF119851	Homo sapiens	PRO1722	162	57
26	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
27	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
28	U47924	Homo sapiens	C8	1112	86
29	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	84	45
30	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	92	65
32	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	qin-induced kinase	206	53
538	AF135022	Homo sapiens	mediator	128	100
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
540	AF016430	Caenorhabditi s elegans	contains similarity to a BR-C/TTK domain	853	39
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308)	408	66
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
543	AF102530	Mus musculus	olfactory receptor F3	327	73
544	Y73431	Homo sapiens	Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84.	386	100
545	AE004833	Pseudomonas	probable TonB-dependent receptor	279	42
-1-	002702	aeruginosa	Human secreted protein, SEQ ID NO: 7874.	264	53
546	G03793	Homo sapiens	A human monocyte-macrophage apolipoprotein	1772	67
547	Y69192	Homo sapiens	B receptor protein.		
548	¥91493	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:166.	176	100
549	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584_2 protein sequence.	1224	94
552	X98330	Homo sapiens	ryanodine receptor 2	24621	99
553	Y42782	Homo sapiens	Human UC Band #331 protein.	684	95
554	AB025258	Mus musculus	granuphilin-a	501	41
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein S239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	183	32
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus musculus	Ca2+ dependent activator protein for secretion	1010	93
561	AF187325	Canis familiaris	melanoma antigen	287	55
562	AJ001981		OXAIL	2512	99
563	Z17238	Homo sapiens Rattus	glutamate receptor subtype delta-1	338	66
564	W30638	norvegicus Homo sapiens	Partial human 7-transmembrane receptor	371	100
			HAPO167 protein.		
565	AC005620	Homo sapiens	R33590_1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63.	1138	78
567	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1002	58
568	AF151043	Homo sapiens	HSPC209	798	100

SEQ	Accession	Species	Description	Smith-	%
ID Č	No.	[ ]		Waterman	Identity
NO:		1		Score	
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
572	AL031177	Homo sapiens	sequence. dJ889M15.3 (novel protein)	735	55
	1			254	45
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	1883	99
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor		
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ ID NO:388.	77	70
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis familiaris	D4 dopamine receptor	64	56
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2 Antigen)	268	85
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
				182	79
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	764	80
589	AF235017	Mus musculus	2P1 protein	704	
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	329	81
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted protein.	110	43
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer polypeptide.	1369	92
593	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein sequence SEQ ID NO:108.	1112	97
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus musculus	COP1 protein	2215	95
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598	AF192499	Mus musculus	putative secreted protein ZSIG37	143	40
599	AF119855	Homo sapiens	PRO1847	236	76
900.	G02872		Human secreted protein, SEO ID NO: 6953.	212	73
601	Y00295	Homo sapiens Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936		diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
605	AB033106	Homo sapiens	Antigen) KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	199
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
610	Y27868	Homo sapiens	HPMBQ32.  Human secreted protein encoded by gene No. 107.	116	62
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
015	1		clone HTDAD22.	1.50	84
614	M87053	Rattus	lens membrane protein	450	04
	M87053 AC004232	Rattus norvegicus Homo sapiens	lens membrane protein FPM315	163	37

PCT/US01/03800

NO:	I			Score	Identity
517	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
518	AJ245621	Homo sapiens	CTL2 protein	2258	99
	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
	D90721	Escherichia	Transmembrane protein dppC	573	90
		coli	Human secretory protein of clone CS752-3.	730	100
	W75858	Homo sapiens	Human secretory protein of clone C3732-3.	733	100
	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	637	83
624	AF034745	Mus musculus	LNXp80		
625	U42580	Paramecium bursaria Chlorella	Pro-rich, IPPPNMSLPLS (3x)	94	46
		virus 1		194	70
626	U79260	Homo sapiens	unknown		50
627	R95913	Homo sapiens	Neural thread protein.	99	100
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
	AD012280	Homo sapiens	DUSP6	414	76
636	AB013382	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
637 638	G02872 M95762	Rattus	GABA transporter	924	89
	000500	norvegicus	Human secreted protein, SEQ ID NO: 7870.	219	60
639 640	G03789 Y01400	Homo sapiens Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAOBK61.	615	62
(42	4 D016002	Homo sapiens	serine/threonine kinase	485	98
643 644	AB015982 Y25806	Homo sapiens	Human secreted protein fragment encoded from	162	46
			gene 23.	<del></del>	<del></del>
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens		233	73
650	G02872	Homo sapiens	100 000	173	78
651	Y32199	Homo sapiens		1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens		57	37
655	L22455	Rattus	mu opioid receptor	116	34
<u> </u>	1	norvegicus	TY CEO ID NO. 7102	110	45
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	459	97
657	G02345	Homo sapiens		291	75
658	W88627	Homo sapiens	HPMBQ32.		
				134	1 65
659	G02832 Y91423	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	134 333	96

SEQ D	Accession No.	Species	Description	Smith- Waterman	% Identity
TO:				Score	<u> </u>
61	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
62	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	375	43
63	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
64	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein (hs6M1-21))	480	55
565	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
567	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KJAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G02332 G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens		275	75
698	Y87280	Homo sapiens		576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens		610	79
701	AF209198	Homo sapiens		2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans		920	51
705	G02882	Homo sapiens		589	98

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
10:	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
06			Tumor necrosis factor receptor 1 death domain	121	95
07	R95326	Homo sapiens	ligand (clone 2DD).	125	39
08	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.		98
09	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	
10	M63577	Saccharomyc es cerevisiae	SFP1	131	59
711	AB026291	Rattus norvegicus	acetoacetyl-CoA synthetase	467	85
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota marmota	olfactory receptor	615	83
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
	AB033062	Homo sapiens	KIAA1236 protein	1380	100
715			Human secreted protein, SEQ ID NO: 4658.	80	73
716	G00577	Homo sapiens	SEO. ID. 37 from WO0034474.	835	99
717	Y96864	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
718 719	AJ243396 U47334	Homo sapiens Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
		<del>                                     </del>	receptor beta4 subunit	1096	100
720	AB020598	Homo sapiens	peptide transporter 3	570	74
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.		100
722	J05046 ·	Homo sapiens	insulin receptor-related receptor	6787	41
723	AF001958	Ambystoma tigrinum	electrogenic Na+ bicarbonate cotransporter, NBC	111	94
724	AF127084	Mus musculus	semaphorin cytoplasmic domain-associated protein 3A	5253	
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus norvegicus	potassium channel	370	100
727	AB029559	Rattus .	BATI	139	35
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila) homolog)	142	68
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor homologue Vanilrep1.	675	99
	1	<del></del>	HSPC264	492	94
732	AF161382	Homo sapiens		3826	99
733 734	AB029033 AE000493	Homo sapiens Escherichia	KIAA1110 protein putative transport protein	592	97
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor (rhodopsin family) protein similar to high-	2173	99
736	AF132599	Homo sapiens	affinity lysophosphatidic acid receptor homolog) RANTES factor of late activated T lymphocytes-	245	56
	1		acetylcholine receptor delta subunit	883	99.
737	X55019	Homo sapiens	acciyichonne receptor dena subunit	1978	100
738	X91906	Homo sapiens		1444	98
739 740	AB026116 D00570	Homo sapiens Mus	organic anion transporter 4 open reading frame (196 AA)	83	24
	1	musculus	Human thyrotropin GPR N-terminal sequence.	118	40
741	W03626	Homo sapiens		614	100
742	U66059	Homo sapiens		2751	99
743	AF119815	Homo sapiens	G-protein-coupled receptor	148	93
744 745	X16663 W67838	Homo sapiens Homo sapiens	Human secreted protein encoded by gene 32	448	95
L	]		clone HLTCJ63.	10014	100
746	W57260	Homo sapiens	Human semaphorin Y.	2414	. 100
747	W21578	Homo sapiens	from plasmid pGCS2232.	968	65
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein sequence SEQ ID NO:76.	622	100
749	AL022238	Homo sapiens	dJ1042K10.5 (novel protein)	314	85
750	G03889	Homo sapiens		391	87

SEQ D NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
51	AB025258	Mus musculus	granuphilin-a	773	41
52	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
53	Y48586	Homo sapiens	Human breast turnour-associated protein 47.	2527	99
54	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus norvegicus	vasopressin receptor	979	68
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsiblefor binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8Bi92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus musculus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 delta	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOV1	1904	91
786	Y91870	Homo sapiens		547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens		745	96
791	Y08639	Homo sapiens		1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
794	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	124	62
795	Y69384	Homo sapiens		119	100
1		Homo sapiens		1358	99

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO:	A PO50240	Tierre coniene	hepatocellular carcinoma-associated antigen 112	1151	99
797	AF258340	Homo sapiens	nepatocentilar carcinoma-associated antigen 112	461	98
798	AF159615	Homo sapiens	FGF receptor activating protein 1	797	99
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	572	92
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.		
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens (human)	CR1 protein.	11963	97
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
		Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
806	G01731			1562	83
807	Z24680	Homo sapiens	garp glycoprotein-associated amino acid transporter	1364	90
808	AF171669	Homo sapiens	LAT2		
809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115 clone HOVBA03.	855	99
	1 11100001	TY		4561	100
811	AF108831	Homo sapiens	K:Cl cotransporter 3	862	100
812	AF092135	Homo sapiens	PTD014	784	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899		
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone gn114 1.	358	100
017	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
817		Homo sapiens	CGI-41 protein	1106	95
818	AF151800			3980	100
819	L00352	Homo sapiens	low density lipoprotein receptor	5832	99
820	X04434	Homo sapiens	IGF-I receptor	572	100
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.		48
822	AF212220	Homo sapiens	TERA	396	
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored protein GPI-122.	4897	99
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Home sapiens	neuronal voltage-gated calcium channel gamma-	1105	100
023	AF030322	Home suprens	2 subunit		
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded from gene 28.	1540	100
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by	541	98
	<del> </del>	<del></del>	gene 24 SEQ ID NO:147.	1625	87
830	X54232	Homo saptens	glypican	2540	100
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein		
832	Y71262 .	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828_1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi s elegans	glycine-rich	85	36
837	AL121889	Homo sapiens	AL023803))	998	75
L		Homo sapiens	plexin-B1/SEP receptor	1580	60
838	AJ011415		1 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1105	67
838 839	AJ011415 W80398	Homo sapiens			
		Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
839 840	W80398 G00862	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.	255 644	97
839 840 841	W80398 G00862 G02650	Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.	255	
839 840	W80398 G00862	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein	255 644	97
839 840 841 842 843	W80398 G00862 G02650 AF036717 Y73446	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	255 644 2629 1089	97 99 100
839 840 841 842 843	W80398 G00862 G02650 AF036717 Y73446 G02872	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_l protein sequence SEQ ID NO:114.  Human secreted protein, SEQ ID NO: 6953.	255 644 2629 1089	97 99 100
839 840 841 842 843	W80398 G00862 G02650 AF036717 Y73446	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.  Human secreted protein, SEQ ID NO: 6953.  CGI-52 protein	255 644 2629 1089	97 99 100

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
10:		ł		Score	
_			to AF038969 (PID:g2827207)		
48	X99886	Homo sapiens	monocyte chemotactic protein-2	160	76
49	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to P34984 (PID:g464305)	963	98
50	AB038237	Homo sapiens	G protein-coupled receptor C5L2	1767	100
51	AF124490	Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
352	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	1189	99
353	AF224741	Homo sapiens	chloride channel protein 7	3748	99
354	X17094	Homo sapiens	furin (AA 1-794)	3550	99
355	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
862	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (	1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by gene 25 SEQ ID NO:305.	1048	98
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	94
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875	Y27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
	W63681	Homo sapiens	Human secreted protein 1.	1652	99
877 878	L27867	Rattus norvegicus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
		Homo sapiens	cathepsin L	209	72
883 884	Y18462 Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
885	AF070661	Homo sapiens		404	100
	Y04315	Homo sapiens		385	100
886	X92744	Homo sapiens		375	100
887 888	Y22496	Homo sapiens		994	94
990	V41202	Homo sapiens		4595	99
889	Y41293			147	63
800	G03714	Homo sapiens Homo sapiens		1012	99
890	AF208856	Homo sapiens		2002	98
891	1100107	i Homo saniens		1953	100
891 892	U29195				
891	U29195 X68149 Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein	537	100
891 892 893 894	X68149 Y94914	Homo sapiens Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	
891 892 893 894	X68149 Y94914 W61630	Homo sapiens Homo sapiens Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.  Clone HNFGW06 of EGFR receptor family.	537 326	63
891 892 893 894 895 896	X68149 Y94914 W61630 M24110	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.  Clone HNFGW06 of EGFR receptor family.  G0S19-2 peptide precursor	537 326 481	63 100
891 892 893 894	X68149 Y94914 W61630	Homo sapiens Homo sapiens Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.  Clone HNFGW06 of EGFR receptor family.  G0S19-2 peptide precursor  imogen 38	537 326	63

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.	-		Waterman Score	Identity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
903		Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
904	Y00261			771	99
905	AF039688	Homo sapiens	antigen NY-CO-3	2544	100
906	AB007836	Homo sapiens	Hic-5		100
907	AB017507	Homo sapiens	Apg12	224	98
908	AK000056	Homo sapiens	unnamed protein product	1537	
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo 1	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta	1319	100
1	114154	TACALLO GAPAGA	protein sequence.		1
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
010	7/10000	77	SEQ ID NO 475 from WO9922243.	1361	100
913	Y19757	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
914	G03172	Homo sapiens		886	90
915	U14971	Homo sapiens	ribosomal protein S9	1204	99
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1963	100
917	AC005525	Homo sapiens	F22162_1		
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
924			down-regulated in gastric cancer	785	78
925	AB039886	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
926	G03368	Homo sapiens	Human breast tumour-associated protein 67.	539	100
927	Y48606	Homo sapiens		668	100
928	Y36151	Homo sapiens	Human secreted protein #23.		100
929	AF110399	Homo sapiens	elongation factor Ts	1666	99
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
930	X30383	familiaris	1400	1	
937	B08906	Homo sapiens	Human secreted protein sequence encoded by	117	44
	1	<del></del>	gene 16 SEQ ID NO:63.	1064	99
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	515	42
939	Y53886	Homo sapiens	designated HSCOP-6.		
940	Y16630	Homo sapiens	(PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens		1289	81
943	AF226046	Homo sapiens		1049	98
944	Y36078	Homo sapiens		667	100
24-	100000	Home series		565	100
945 946	M22877 W67869	Homo sapiens Homo sapiens		551	93
947	W67859	Homo sapiens	clone HHGDB72.	283	100
177	1,0,035	120:20 544:015	clone HBMCL41.		
<u> </u>	W85726	Homo sapiens		789	100
948		, omp.o			
948	AJ242015	Homo sapiens	eMDC II protein	4236	100

EQ D	Accession No.	Species	Description	Smith- Waterman Score	% Identity
10:	17110645	TTo-o coniona	candidate tumor suppressor p33 ING1 homolog	1314	100
51	AF110645 Y36111	Homo sapiens Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
~			ID NO. 496.		100
53	AB012109	Homo sapiens	APC10	990	100
54	AF246221	Homo sapiens	transmembrane protein BRI	1405	
55	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
56	W74726	Homo sapiens	Human secreted protein fg949_3.	1879	100
57	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
58	AJ222967	Homo sapiens	cystinosin	1920	100
59	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein sequence SEQ ID NO:110.	587	100
60	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
61	AF151855	Homo sapiens	CGl-97 protein	1214	96
62	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
63	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
64	AF078859	Homo sapiens	PTD004	2089	100
65	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc# AF030433	1466	100
66	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
67	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
68	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
69	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
770	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972 ·	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide gated cation channel hHCN4	6295	100
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
977 978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP80	3348	99
979	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
000	4.5110202	Homo sapiens	neuroendocrine-specific protein-like protein I	1170	99
980	AF119297 AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
981 982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412- encoded protein.	1553	99
002	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
983	AB026125	Homo sapiens	ART-4	2160	100
984 985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	172	70
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	ar transfer to the state of the	712	100
988	AF153450	Manduca	juvenile hormone esterase binding protein	226	32
000	G03607	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
989 990	G03697 AF204159	Homo sapiens		1486	100
001	002061	Homo sapiens		558	99
991	G02061	Caenorhabditi		327	40
992	AL031266	s elegans	<u> </u>	4730	99
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	141	77
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	5811	99
995 996	AF133845 AF117756	Homo sapiens	thyroid hormone receptor-associated protein	4999	100
ĺ			complex component TRAP150	284	93
	W62066	Homo sapiens	Human stem cell antigen 2.	725	100
997		Homo sapiens	Human secreted protein sequence SEQ ID	123	1 *00
998	Y87173		NO:212.	1651	- 00
	Y87173 Y13379 Y95008	Homo sapien	Amino acid sequence of protein PRO263.	1654 676	99

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
003	W73420	Homo sapiens	Human secreted protein encoded by Gene No. 24.	2150	100
004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
005	M23323	Homo sapiens	membrane protein	642	100
006	X63745	Homo sapiens	KDEL receptor	326	98
007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
800	AB032918	Hylobates moloch	dopamine receptor D4	92	35
009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
010	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
014	AF288092	Naegleria gruberi	haem lyase	114	37
015	AB045292	Homo sapiens	M83 protein	3867	99
016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
020	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
023	AE000660	Homo sapiens	hADV36S1	573	100
024	AF132965	Homo sapiens	CGI-31 protein	1550	100
025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMP1	806	100
030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1032	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID	639	99
			NO:1.  Human secreted protein clone yd178_1 protein	752	93
1034	Y73473	Homo sapiens	sequence SEQ ID NO:168.  Human gene 48-encoded protein fragment, SEQ	96	90
	Y86468	Homo sapiens	ID NO:383.	698	100
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	3699	99
1037	AJ242832	Homo sapiens	calpain	2574	100
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	1310	100
1039 1040 ·	AJ242730 AF169968	Homo sapiens Mus	polyhomeotic 2  DNA binding protein DESRT	1453	80
1041	V52562	musculus	permability increasing protein	383	29
1041	X52563	Bos taurus	Human secreted protein, SEQ ID NO: 4449.	75	50
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.  Human secreted protein, SEQ ID NO: 6613.	60	53
1043	G02532	Homo sapiens		1850	100
1044	M94582	Homo sapiens	interleukin 8 receptor B		50
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	30
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049		Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
	1	Homo sapiens	liver-specific transporter	2820	100

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:		l i		Score	<u></u>
051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	770	85
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082	L13802	Homo sapiens	ribosmal protein small subunit	499	80
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42 clone HSXB125.	143	81
1084	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	+83	51
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	88	43
1086	AF090942	Homo sapiens	PRO0657	124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEO ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens		490	70
1099	G04063	Homo sapiens		83	35
1100	Y02693	Homo sapiens		149	59
1100	102073	Troung achiens	clone HTDAD22.	1	1 - 7

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO: ]				Score	72
101	AF119851	Homo sapiens	PRO1722	183 207	62
102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	91	52
103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	128	69
104	X74856	Mus musculus	ribosomal protein L28		
105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
109	AF201951	Horno sapiens	high affinity immunoglobulin epsilon receptor beta subunit	738	94
110	AF111108	Mus musculus	transient receptor potential 2	223	79
111	AF119900	Homo sapiens	PRO2822	144	59
112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by gene 121.	164	63
115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
116	X51394	Xenopus laevis	APEG precursor protein	130	40
117	M27926	Homo sapiens	neutral protease large subunit	442	65
117	M27826	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
118	G03371		Human secreted protein, SEQ ID NO: 7432.  Human secreted protein, SEQ ID NO: 7683.	491	97
1119	G03602	Homo sapiens	Extended human secreted protein sequence, SEQ	244	97
1120	Y35906	Homo sapiens	ID NO. 155.	122	65
1121	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.		90
122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469	Homo sapiens	Human secreted protein from clone CW795_2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by gene 49.SEQ ID NO:170.	542	100
1134	AB017908 ·	Homo sapiens	4F2 light chain	2399	93
1134	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis familiaris	D4 dopamine receptor	89	48
	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1144	ı		שלים שלים שלים שלים שלים שלים שלים שלים		-+
	V00062	Home caniers	rah-related GTP-hinding protein	1 398	196
1145	X99962	Homo sapiens	rab-related GTP-binding protein Human secreted protein, SEO ID NO: 7888.	398 168	96 79
1145 1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1145			Human secreted protein, SEQ ID NO: 7888.  Human secreted protein, SEQ ID NO: 7793.		

SEQ ID	Accession Species Descri No.		Description	Smith- Waterman Score	% Identity
NO:		s elegans	cerevisiae zinc resistance protein	500,0	<del> </del>
1160	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1150 1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181.	80
		Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1152	G03798			95	41
1153	X88799	Oryza sativa	DNA binding protein	155	96
1154	D85245	Homo sapiens	TR3beta 62		87
1155	R74272	Homo sapiens	Tumour suppressor protein, p53.	341	
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	99	41
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
		Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1170	Y45274			347	96
1171	D64154	Homo sapiens	Mr 110,000 antigen	311	67
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	60	52
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.		
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus norvegicus	tRNA selenocysteine associated protein	249	62
1182	AF161524	Homo sapiens	HSPC176	138	90
	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1183		Homo sapiens	Human secreted protein encoded by gene 22	107	71
1184	Y02671		clone HMSJW18.		
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	118	46
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.		37
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens		324	76
1191	AF117755	Homo sapiens	complex component TRAP230	187	70
1192	Y70455	Homo sapiens	5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens		192	76
1195	W29661	Homo sapiens		2001	98
1196	Y14104	Homo sapiens		239	69
1197	X61972	Homo sapiens		149	90
1197	G00534	Homo sapiens		145	51
		Homo sapiens		1089	89
1199	Y86260	Homo sapiens	NO:175.	154	57

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO:		YY	Human secreted protein, SEQ ID NO: 4919.	404	50
1201	G00838			202	49
1202	M27826	Homo sapiens	neutral protease large subunit	265	61
1203	¥73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.		
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1213	AF090942	Homo sapiens	PRO0657	124	70
	Y14427	Homo sapiens	Human secreted protein encoded by gene 17	99	77
1215			clone HSIEA14.	173	57
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	1173	100
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	454	78
1218	J00194	Homo sapiens	hla-dr antigen alpha chain		92
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-2) polypeptide.	725	100
1221	W96745	Homo sapiens	High affinity immunoglobulin E receptor-like protein (IGERB).	650	98
1222	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 160.	135	31
1223	Y00278	Homo sapiens	Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226	G01733	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	YIPIB	801	63
	1	Homo sapiens	soluble adenylyl cyclase	275	100
1230	AF176813		organic cation transporter	1704	100
1231	X98333	Homo sapiens	Human secreted protein encoded by gene 77	212	53
1232	W74955	Homo sapiens	clone HOEAS24.	526	100
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1-protein sequence SEQ ID NO:86.		
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens		697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1800	V00710		Human membrane-associated protein Zsig24.	709	97
1241	Y92710	Homo sapiens			88
1242 1243	Y95002 Y44905	Homo sapiens Homo sapiens	Human potassium channel molecule ERG-LP2	325	100
L	_1		partial protein.	+ = 1 :	97
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	
1245	Y53629	Homo sapiens	BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1247	Y35911	Homo sapiens		168	39

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
			ID NO. 160.		
248	AF072509	Rattus norvegicus	glutamate receptor interacting protein 2	559	90
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
250	B08974	Homo sapiens	Human secreted protein sequence encoded by gene 27 SEQ ID NO:131.	1087	97
1251	L15313	Caenorhabditi s elegans	putative	858	59
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate reading frame protein.	278	75
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-l	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded from gene 26.	81	94
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	986	100
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor sequence.	288	71
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80:
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ ID NO:2.	723	93
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus norvegicus	phosphatidylinositol 5-phosphate 4-kinase	859	95
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus musculus	LMBR2	552	76
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras GTPase-activating protein p135 SynGAP)	820	98
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthethase (acetate-coA ligase))	1280	100
1274	AF064748	Mus musculus	\$3-12	3523	61
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	-Y30715	Homo sapiens	Amino acid sequence of a human secreted protein.	643	90
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus cuniculus	aorta CNG channel (rACNG)	267	85
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1283	AF117814	Mus musculus	odd-skipped related 1 protein	357	98
1284	U87318	Xenopus laevis	NaDC-2	535	60
1285	AF061346	Mus musculus	Edp1 protein	452	68
1286	AB030182	Mus musculus	contains transmembrane (TM) region	582	68
1287	A13595	synthetic construct	immunosuppresive protein PP15	185	97
1288	AF254411	Homo sapiens		837	100
1289	AF084205	Rattus	serine/threonine protein kinase TAO1	319	98
1		norvegicus			]

EQ D	Accession No.	Species	Description	Smith- Waterman Score	% Identity
10:	A 77029662	Tions sociens	membrane associated guanylate kinase 2	523	100
290	AF038563	Homo sapiens	double-stranded RNA specific adenosine	468	100
291	AF034837	Homo sapiens	deaminase		
292	M15888	Bos taurus	endozepine-related protein precursor	937	87
293	AB010692	Arabidopsis thaliana	ATP-dependent RNA helicase-like protein	636	45
294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
295	W67828	Homo sapiens	Human secreted protein encoded by gene 22 clone HFEAF41.	504	98
296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1202	Mageon	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
302	M77693		Human secreted protein, SEQ ID NO: 5412.	254	69
303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 3412.  Human secreted protein, SEQ ID NO: 5572.	747	99
304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 3372.	602	98
305	AF148509	Homo sapiens	alpha 1,2-mannosidase	333	98
306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.		98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	332	
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
1309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
			sperm antigen	2607	98
1317 1318	AB041533 U19617	Homo sapiens Mus	Elf-1	806	92
1319	U82598	musculus Escherichia	ferric enterobactin transport protein	768	100
1320	D90892	Escherichia coli	SORBITOL-6-PHOSPHATE 2- DEHYDROGENASE (EC 1.1.1.140) (GLUCITOL-6-PHOSPHATE DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).	709	100
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41 clone HPBCJ74.	601	92
1322	AJ276101	Homo sapiens		466	93
1323	AJ276101	Homo sapiens	GPRC5B protein	504	97
1324	Y58628	Homo sapiens		1584	100
1325	U91561	Rattus norvegicus	pyridoxine 5'-phosphate oxidase	1277	89
1326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
1327	Y32206	Homo sapiens		1531	90
1328	AF151048	Homo sapiens		657	85
	Y10530	Homo sapiens		1645	100
1329		Homo sapiens		4314	99
1330 1331	AF180681 AF111856	Homo sapiens	1	3591	99
1331			( (VaC != 11)	1	
1332	Y13583	Homo sapiens		2171	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

## TABLE 3

CDO ID	DEC TE	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	noa	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914		acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		j	,	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide	sequence	/=possible nucleotide deletion, \=possible
İ	,	1		sequence		nucleotide insertion
	1551	<del>                                     </del>	2	337	<del>                                     </del>	TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
1	1351	Α	2	337	ļ <b>"</b>	HWPQAPHRA***GLLPPRWLGHGLPGGPAAP
	1	1			İ	WAASQWVDGVAGRLPGPAWSWHASGAAPA
			1			OPGPL*LLVPGSSGLPDPRDP
		<del> </del>	-	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL
2	1352	A	2.7	100	300	OIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT
	1					GTLIHGWWGCKVVEPLGKTVWQIPK
	10.55	<del>  </del>	1.40	<u> </u>	314	HASAHASVVLKDNSELEQQLGATGAYRARA
3	1353	A	40	3	314	LELEAEVAEMRQMLQLEHPFVNGADKLRPD
	ļ	1	1	}	1	SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R
	1	1	1			SCNYTLALILFL
	1	<del>                                     </del>	1	<del></del>	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR
4	1354	A	74	2	292	QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT
	1	1	}		}	VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH
			1			VLPLP
<u></u>	L	1	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP
5	1355	A	/8	114	930	NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA
		1		į.	ļ	FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI
		į.			1	AGMLGAVISGIWLDRSKTYKETTLVVYIMDT
						GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF
<u> </u>	1			İ	İ	MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA
	1	1		1		OVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG
		1		1		AALTAFIKADLRRQKANKETLEN
L		<del></del>	<del></del>	+	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A
6	1356	A	81	97	370	YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE
}	-	1			ľ	*CLFQEMGLSLQWLYSARGDFFRATSRL
	1	٠	1	<del></del>	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP
7	1357	A	93	2	8/2	ATALADNKPVAPDRRISGHVGIIFSMSYLESK
	İ	1		1	1	GLLATASEDRSVRIWKGGDLRVPGGRVQNIG
1				l	1	HCFGHSARVWQVKLLENYLISAGEDCVCLV
L						TOTOTISATE WAS A REDERITE TOTOTO OF COLUMN

OTO ID	1000	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	{	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	"""	Ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
COLIDO		ļ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	]	}	}	peptide	] -	/=possible nucleotide deletion, \-possible
		]		sequence		nucleotide insertion
						WSHEGEILQAFRGHQGRGIRAIAAHERQAWV
			1		}	ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ
		{	-	1		VP**ARYTQGCDSGWLLATAGSD*YRGPVSL
		ļ			1	*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL
	l	{			[	RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS
						WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
	Ì			ļ		LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
	1	}	1	}		LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV
			<u> </u>		100	QCLGFVDSDSRKMVSTLT
9	1359	A	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN KSSEFNEGPERERMDV
	1	<del></del>	1.00	<del> </del>	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY
10	1360	A	123	2	1249	FEEVORLRFEVHDISSNHNGLKEADFLGGME
l	1	1		1	1	CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA
(	-	1				EELSGNDDYVELAFNARKLDDKDFFSKSDPF
		1				LEIFRMNDDATQQLVHRTEVVMNNLSPAWK
ł		1			1	SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK
		-	1			HDFIGEFTSTFKEMRGAMEGKQVQWECINPK
		1 *	1	1	1 .	YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI
		}		}		MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP
1	1					YQPNEYLKALVAVGEICQDYDSDKMFPAFGF
ļ	1	1	}	1		GARIPPEYTDSHDFAINFNEDNPECAGIQGVV
				İ		EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR
	1	1.		1		SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP
	- "	1	-		1 _	DNPGGHFV
11	1361	A	147	614	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT
		1	ı			NTTSSRPASSRG\TLSSSSSSSSLTKDALPSSL
		1		ļ	ļ	KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI
		1	1	ì	ł	KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS
		1		1		DERVSMGTSSRKPTNSSSSLGALKMSATS\*G
i				1		SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS
						TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM
		-				VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA
		ĺ	1			ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI
		-[	1			NLLTMNYY
<u> </u>			1240	52F	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ
13	1363	A	249	535	103	MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS
	}				1	TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI
1		ł	1	1		DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC
1		ì	ļ			QVLGTPKKVSTLVPKLL
14	1264	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH
14	1364	A	234	3/2	201	FFFLDICYSSVTAODAAEFPVS*KPILVWGYIT
1	1			]		*SFFFIFSWGTNGCLLSAITYACYAAICHPLLS
1			1	1		TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM
13	1305	1^	1 '	1	1	AIEFLLECDQNIT\KLICENT*KNIAKNI*KRRV
		1		1		TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI
			1			KOTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK
10	1300	1		1.5	1	SKDVFSKPVNIFWALEESVLGVKARQPKPFFA
1		- }		· l	1	AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV
1			i	1		GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	-	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW
1 *′	1507	1 "	1	1		VSSLAMKEMLTKTTM
1						
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL ELLTSGDPPALASQSAGITGMSHCARPKGHFG

- AN		T 3.2	Lopo	Dan 25-12-3	Dund: 4-4 3	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NU: 01 nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	acince	{	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
UCILCO			1 7."	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł	1	Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	ł		peptide		/=possible nucleotide deletion, \=possible
		1	1	sequence	ļ	nucleotide insertion
			<del>                                     </del>			IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF
	1	1	1	<b>\</b>		KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR
	1	1	1		1	WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK
	.	1		Į.	ļ	GQPHQSTNALLRRCVR*RYHLS\TVETAGLP*
	}	l		1	ł	KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q
	ļ		ì	ł	}	NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK
1	1	ļ	ļ			TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR
ļ		L			l	ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS
İ		1	1	i	i	WICRLRPLLWRAVREYLSKLKNAELSFDPGV
ł	ł	1	1	Ì	ł	SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC
1	1	1	1	ļ	İ	KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA
l			1	<u>]</u>		AV*NKPRHLLSHIWKDVQNILLK
20	1370	A	304	1	1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA
		1	1		ł	LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA
	1	1	1	1		GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG
1	ł	}		1	1	GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP
ļ	1 :	1	1	} .	}	CPHPPGFRLWMSPNQKPPTENPGVMGRVWR
1				ì		LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA
ļ	Ì					PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF
	1	1	-		(	MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP
ł	1 .	i	1	1		DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR
		1				CRALPGRICSAPAAGLRRARPRISESRRGNSP
,	1	1	}		}	PASPAAASARCPSWGPSCPARPPSRPAAGTEP
}		}	1			AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP
			ì		į	VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW
1 .	1	1	1			ALVRSRGG
21	1371	A	326	799	1587	GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC
2.	13/1	1	} 320	1 ""	1007	E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP
1	1	1	1			HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP
ì		1		ì	1	LRHVRLFSAGAPRGAATPCPPALLHGPAWPP
1	1	1	1		}	ARPMFRGHPPVRPLGPWGKVAAGPRALCLA
1	ļ	1	1	1	ĺ	GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL
1	1	1	1			QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA
)	-			}		AAQAEPGADPEPEDKDQAAESRPAGAMSLSA
1		ŀ	Į	1	}	QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP
			1		1	GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL
	1	1		ţ		PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL
1			1			SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP
1		1		ļ	i	APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP
				1		GTVVSP
23	1373	A	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES
1 .		1	1	1	1	LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL
		1	1			NNEKRKMKKRKEEKKKCRERMQRRSKWRR
		-		1		EEKKE*RREE\EERKKEKEDRKERRKETSPRG
			<del> </del>	<del> </del>	1000	SRRLLRD
24	1374	Α	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP
1			<del></del>	<del> </del>	100	WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ
25	1375	Α	384	373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD
1		1	1			
				1252	+	KKINLNLKPHTKLTPNIKKN
26	1376	Α	397	383	165	EVKNTNPFIFSGTNLTIWIRSI*RKSDEINQRTK
1	1	1		1	1	*MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI
	1		100	102	1200	KIPANF KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK
27	1377	A	406	103	380	IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI
	1			1		LVNKIEDLNKWRNYLLSWIGRRNIINTMT
L						TALVIED TIVE AVIA A TES MICKGAILIM I MI

050 15	COPO TO	34-	LOPO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	noa	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
neuce	dance		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciicc		}	124	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		i	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		}	}	peptide		/=possible nucleotide deletion, \=possible
	j	)		sequence		nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL
				j ·		ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF
	ł	1		l		YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI
		1	1	İ	I	TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA
	1	i	ľ	1.	Í	NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK
	1		1		ł	VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID
			İ	l		PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS
		1	1	ì	į	SWDYRYAPPRP\ANF\*FLVETGFYYVAQAGL
	L	L				KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI
		1		1		CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV
	Į	ì	1	l.		EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK*
		[			1	KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK
	1500	+	<del> </del>	105	471	IFAN VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK
32	1382	Α	474	125	471	EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ
					1	KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT
	1	1	į.	1	1	ILRETDRIHKTTYDVISLI
22	1383	A	488	1825	+2	KSACSFICSEEQPASPSPLKPGTYASET\RPRDP
33	1363	I A	400	1023	1	HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT
}	1		}		1	PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR
		1	1			APGGSAGS*\GLPSAGGSRGRKGWRAAGRQP
}		1		1	1	STR*GRPGRHGGRGE*AGHPEPRQSALQSAG
1		[	1			L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P
Ì		1		1		PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP
ł	1	1		ľ		SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG
]		1				AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA
]		1		1		SPQTAAGAGSPVQWALSRATG*TGETGSWC
1		1	}	1		AGGTHQATHLTAAWVCPPTWSVRPGGSGPA
1		1	1	1	1	AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP
i		1				SPASSEVALSSGSCWPDQAPGPARGSPPAPLA
	·			Ì		PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL
		1	)	}		SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP
		1		i		L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS*
	1	1	\ ,			GGRSPAGTGHLGAQTVASPH*GHWPTALSCL
	1	Ì	1			WASASPPGPEAPPQTGACIGTNCRYRAASAR
1		}	İ		1	RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
L		<del> </del>	105	122	+	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E
34	1384	A	497	422	2	AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA
	1			.		RLS\PPLASCGGRGPPGGAACATCAPPAGPAR
}		1				SSRCRRSPPE*GPR*PSRPARPSPGSAASRRQ
1		1		1	1	KLTPCRCQFRGLCA
25	1205	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN
35	1385	^	209	130	17/3	LTELVVAVTDENIVGLFAALLAERRVLLTAS
i		1		1		KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH
1		ĺ		{	4	LLDYC*CPPLPRT
36	1386	A	512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA
30	1300	^	312	1	1031	FLGLAAGGOTLCPAGELPGHARAQASGAPGS
		ŀ		1	1	VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL
.[		ł				GHTRPARPRPV\PFAPAVPQEPGGQGHGAA/P
1	}	1		1	1	PATGHSAPRGCPPARAAPTGSATPAPPPAACA
			1	1		AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT
1					j	PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA
1	-	-				GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP
1	'	-		1		VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR
}	1	-	Ì		1	AAVARRLRSWNACGLSRVAGRSSASYPGRE
L	l					

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	ucnec		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
detice			1 7.7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			]	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Soque	/=possible nucleotide deletion, \=possible
		i		sequence		nucleotide Insertion
		<del></del>	<del> </del>	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
				j	} .	GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
•		ĺ				VSAAPOSPRTRCPRGCAAAAGLCVLAAAGAS
		İ	ļ	}		HGA\GLPGVRVHTQRVHIH*GAG/GCQTPRPR
			l	i		LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
·	ļ	İ	1	•		ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
	]		1	ļ	ļ	NHQAVGLEASGALQAGHRDELPTMVQLLDH
		1				SPDYPLKGRPHAP
37	1387	Α	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
	]		j			WEASPEMQSKCHQKGKNNQTECFNHVRFLQ
	}		1			RLNSTHLYACGTHAFQPLCAAIDAEAFTLPTS
1		1		ļ		FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE
İ		I	1	1		FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
1		ì				AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
	1	İ				SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH
		i				\DAEFVFSVLVRESKASAVGDDDKVYYFFTE
	}	)	_			RATEKESGSFTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
		ŀ	ļ			TSEEGKEG*EVPV*LPVSPPLPRPLQKMLDYL
	ĺ	1	1			KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
1				1	ĺ	QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT
	İ		1			INIIICTVDYLLRLQESI
39	1389	Α	767	ſſ	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
Ì	ļ	1	ļ	1		RNLSVDGKNVDMAGFIANNGTREGCAARRN
	}	Į		l	1	FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
ì	<b>\</b>	1				KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
		1				VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL
	1	i				RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
		1	1			ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
1	1	}	1			PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
1	1	ł		1		LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
		Ì				TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA
	}	1		1		SEDKVSVRRGFRGCMQVRGGCGGRGEACPS
	ļ	<u> </u>		<del></del>	-	QAAPRL
40	1390	A	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV
		1		1		PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL
	]	1	1			RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
	1.000	1	1035	<del> </del>	105	VNKSVVLVQVTIM SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
41	1391	A	835	7	195	SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT
40	1200	+	-	+,	415	
42	1392	A	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW
1	1		1	1		OGIGTVASNLSQL*TLNAPFPELLLFRSLARTG
1	1		1	i	ļ	FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
	1		1			
1-2	1.200	+	1005	1250	102	KACFPTNIVTL PALSPAPVPOKKGSPLPLDPCLGPSSWLLSVG
43	1393	A	845	358	92	LGWPRL*PRRGPGDPGSLPATTPLLTPPHTLLP
1		1		1	1	I —
1	1.224	<del> </del>	1063	162	1,	QRPMLPPSHAGLARPPPPEPISVP LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK
44	1394	A	853	452	1	
1		1	1	1		SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS
}	1	1		1	}	PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
1	1	1		1	1	QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
1.	1	1,-	1001	1 270	122	PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQTV\$LNTWDTAG
					1	QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
	1	1			<del>   </del>	WLSMSMGK
	+	<del></del>				
46	1396	A	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
46	1396	A	900	1	. 366	EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR VFLFHQLNIT**CLHFFTMTTFIAIPFSFLFLGR

						Amino acid sequence (A=Alanine C=Cysteine,
SEQID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Aianine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		amino acid	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	] .	residue of	Sedirence	/=possible nucleotide deletion, \=possible
		l		peptide		nucleotide insertion
	l			sequence		D/KSLAMLPRLVSNSWPQVILPP
					<u> </u>	QLQNLASRGCL*SQLLRRLRRENRLNPGGGG
47	1397	Α	944	162	2	CSEIAP\CTPAWVTQRDFFRKKK
	1	·	I			HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
48	1398	A	963	216	308	PRKRESWWGERLP/PRGFPPAAEDAPAPGWK
49	1399	A	967	466	1	GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP
	Į.	1	-	١.		RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E
	1	ļ	1		1	GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG
	<b>j</b>	)	1	1	į	AAAAQP**TPRCPAALRAGAHIGRVGRPY
	1	1		l		EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA
50	1400	Α	973	45	421	EKCIQALDVFVFCYIDHSSACLMSCD-EDQA
			1	1	}	LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC
	ł	1	1	1	1	ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY
	1	.l	1	1	ì	VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE
	1	1	1		1	E PER PER PER PER PER PER PER PER PER PE
51	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG
••	1	1				PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA
	ł	1	ı	1 .		WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE
	i	1	Ī	}	1	DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG
				ì		WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS
	1	1	İ	}	ł	ESHAASNDPRNFVPNKMWKGLVKRNASVET
	j	}		1	1	VDNKTSEDVTMAAASPVTLTKGTSAAHLNS
	1	1		ļ	İ	MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT
i	- {	ı	1			AVASSTTAASITTAASSMTVASSAPTTAASST
į	1	l	ì	ł		TVASIAPTTAASSMTAASSTPMTLALPAPTST
	1	1	1			STGRTPSTTATGHPSLSTALAQVPKSSALPRT
ĺ	1					ATLATLATRAQTVATTANTSSPMSTRPSPSKH
	- {	1	- 1	1		MPSDTAASPVPPMRPQAQGPISQVSVDQPVV
		- 1	1	1		NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR
	į.		<b>\</b>		1	EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT
ì	l l		1	1		QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS
1			1			SGGTKMPATDSCQPSTQGQYMV/DHH*APHP
1	l l		1	1		GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL
(	1 -	- (			1	*ELQEEGLHPGGLLNQRDVCGLRNVRGAGA
ì		1				WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC
· ·	İ	ı	1	ļ		QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF
52	1402	1	1	1		*SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR
	1	1	1	1		RYLH*SLRLNGEOLKTFPLRSGMR*G/CAILPL
	- }	j	- 1			VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE
	1 :	- 1	1	1		EK*FS*FVGDMNTCVENKKESKKLLE
-	1402	$+_{A}$	1011	$\frac{1}{1}$	630	PEVIOOSAYDSKADIWSLGITAIELAKGEPPNS
53	1403	A	1011	] ^	1	DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV
	-	1	1	1		*LMLA*TKDPSI\RPTAKELLKHKFIVKNSKKT
		1				SYLTELIDRFKRWKAEGHSDDESDSEGSDSES
1	1	l l	1		1	TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ
		1	-		1	DLVQTLSCLSMIITPAFAELKQQDENNASRNQ
	1		- 1			AIEELEKSIAVAEAAGPG
<u></u>				<del></del>	1222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN
54	1404	A	1016	1	222	TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF
ŀ	- 1		1	1	Ì	PISGSGARI
L					1266	HASVDGDEGSDDVYYYYTPAILRELQALNTA
55	1405	A	1033	3	366	EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL
		- 1	ì	1		HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT
		1	1	ļ	-	QKRAA\LYTWHVLEQLEILRQINQQSHGPG
		1				QKKAALLI WIYLEQLEILAQIIQQSIOI O
56	1406	A	1044	5	429	SVLTLQTRSPSKPLS\RKLMDWEVVSRNSISE
				1		DRLETQSRASRSPPVTPNQSQETPVDGKPLAI
		ļ	l	[		PPNQSQKNIRYHIHYLHLQYYLDRHISATLPII SSSGIPTPIAVITDALTDLVELILGQPCSEESGR
		- 1	ı	1	i	NSSGIPTPIAVITDALTDLVELILGQPCSEESUR
1	ſ		1	1	l l	APGTLFLLAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino	O=Glutamine, N=Asparagine, r=1 toline,
uence		Ì	914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
				amino acid	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ì	ł	•	Sequence	/=possible nucleotide deletion, \=possible
		l	1	peptide		nucleotide insertion
	1407	A	1050	sequence	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
57	1407	A	1030	**	450	MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV
	}			i		TCLRRREKDGVIVDVLSDTASNHNGFPVEEH
	Ì		İ	Ì		ADDTHPARLQGPTLRSQPMGPLKHKAFEERA
,		l	ľ	Į.	ļ	NLGLVORRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL
50	1400	1	1000			PCLGCPTXATCRLYQTTVAVVF
59	1409	TA -	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF
J)	1 100	1		[ ]	•	KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC
	1			1		SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR
	1	1	1	1		HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH
	1					TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL
		1	1		i .	LLLAVQQSCLADHLLTASWGGK/DPIPTKALG
		L				EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCOPFHLVMRDLLQLGQDIPQGCHY
		1		Į.		LEENHIJHRDIAARNCLLSCAAPTRAATIGDF GMARYTYRTRYYOLGDRAL/LPRKWMPPEAL
		١.		1		LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR
				[ -		TN LEGIFIANIDS WIFGVELWEIFSEG IMP IT GR
		<del> </del>	1.000	<del>                                     </del>	950	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER
62	1412	Α	1080	1	859	ANLMHMMKLSIKVLLQSALSLGRSLDADHA
	ŀ					PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF
	ļ	1	İ	}	1	GPLELVEKLCPEASDIATSVRNLPELKTAVGR
ĺ.		1				GRAWLYLALMOKKLADYLKVLIDNKHLLSE
	1	1	1			FYEPEALMMEEEGMVIVGLLVGLNVLDANL\
1					]	CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE
1	1	1		ł	1	HERITDVLDQKNYVEELNRHLSCTVGDLQTK
l			}	Į	· I	IDGLEKTNSKLQERVSAATDRICSLQEEQQQL
		1				REONELIR
63	1413	A	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK
"	1	1	1	l	-	HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI
	ļ			1		HTGEKPYTCGECGKTFRQSANLYAHKKIHTG
1	l l			-		EKPYTCGDCGKTFRQSANLYAHKKIHTG\EKP
j	ļ	Ì	}		İ	YKCKECGKAFKSYYSILKHKRTHTRGMSYEG
1	ł	1			-	DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE
}		<u> </u>				KAFNHTSICCRHKKN
64	1414	A	1084	946-	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA
		1			1	SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD
1	1	1	i	ł	1	LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW
1		1		i	ĺ	TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS
	1	1		1		FFSWLTTGLTTQQRTAIE\NATVAFF\LQCI\SC
1		ļ		1		HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI
1		-	1	1	1	S/GFIR/RLFLQLMLEDEKVTMFLQSPCPLYKG
1		1		1	(	RINATSHVIQHPMYGAGHKFRTLHLPVSTTL
		1				SDVLDRVSDTPSITAKLISKQKDDKKKK
1	1415	+	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA
65	1415	A	1087	103	327	LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA
			,	ļ	}	SVALHKLSNALV
-	1416	A	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ
66	1410	^	1033	١	1.55	PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD
1	1	1	ļ		1	TLPVAAAFTETVNAYFKGADPSKCIVKITGE
	1	ì	ĺ		1	MVLSFPAGITRHFANNPSPAALTFRVINFSRLE
1		1				HVLPNPQLLCCDNTQNDANTK\EFWVNMPNI
1		1.	1		1	MTHLK
		1			<del></del>	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA
67	1/17	A	1008	1.57	356	LKLISLGFIIGVSV VGNLLISILLVKDKI LHKA
67	1417	A	1098	57	356	PYYFLLDLCCSDILRSAICFPFVFNSVKNGST WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	İ	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l	ļ		peptide	<b>\</b>	/=possible nucleotide deletion, \=possible
	İ	İ	l	sequence	l	nucleotide insertion
			1			RYL
68	1418	Α	1106	1	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
		}	1		j .	YEREGMQDWKTASGQSEEATQQSSQKPQPH
					]	YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF
		1		1		PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
	1	i	1			HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA
	ł	1	1			RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
	į.	}	1	Į.	}	SPAALAPRAARGGSRAAALAGAEAEEPLRTL
	ì	1	ľ	ļ		APRPTRAAAPPPPPPPPPPPPLPPGAPPPPVRCVSR
	}	}	<b>\</b>	ļ	j	RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
,	Ì		]	ļ	}	APALQIRKGTSSGLPGRGGGSGPGNNLSSVA
	[	l .		1	İ	GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
		1	İ		ļ	SVIEGVSDQVLVAVVVSFALIATLVYALFRNV
	1	1	1	1		HQNIHPENQELVRVLREQLQTEQDAPAATRQ OFYTDMYCPICLHOASFPVETNCGHLFCGSLT
	1	1	1	1	{	1
	114.0	<del>   </del>	1100	12	166	PNSIW FDTARLHEFGTSITQIFAVDNREDLQKWMEA
69	1419	Α	1107	2	466	FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF
		1	1	{		LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE
		i	İ	1	:	•
	1			1	Ì	TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
	1100	<del>                                     </del>	1,,,,-	1.000		RKRY/SYPASEPLHDEKGKKRQAPLPPSDK ALRRLHYVRATKV\FLSFRRPFWREEHIEGGH
70	1420	A	1111	698	23	
		1	1		1	SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
	ŀ		1	į.	1	AFAGLSREEALRLALDDVAALHGPVVRQLW DGTGVVKRWAEDQHSQGGFVVQPPALWQT
	1	1	İ		1	EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
	1	1	l		1	KSALRAAIKINSRKGPASDTASPEGHASDMEG
}		1	)	1	Ì	QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL
	1	1		j	}	ONTTHTRTSH
71	1421	A	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
	1421	^	1119	*	363	PPGPPEOAGLSOFHLEPETQNPETTEEIQSS\LQ
	1	1	1	-	i	QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL
ł	1	1	1	}		EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	+1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
1 12	1422	1	112/	1 *	] 300	GOKVAIKIVNREKLSESVLMKVEREIAIL\RLI
		.			į	EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
					į	QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG
1						GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
		1		ł		HSYSICHRDLKPENLLLDEKNNIRIADFGMAS
1		1			1	LQVGDSLLETSCGSPHYACPEVIKGEKYDGR
1				ļ	Ì	RADMWSCGVILFALLVGALPFDDDNLRQLLE
1		1				KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
-		1		[		LSLEQIQKHPWYLGGNFIS
73	1423	A	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV
		1		1		FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE
1	1	1		1		MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
-	1	1	1	1	ļ	SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
1	1				ľ	LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
		1		1		GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
i	1		1	1		TAGLNVAAEGARARDMPAQAWDLVERMKN
		1	1		1	SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
	ļ	1	1	1	1	HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT
1	1	1				DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG
		1	1		1	VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
1	1		1	1	1	EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
1	}		1	1		AADPAPVHTTAHPKGA
75	1425	A	1147	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine.
		<u> </u>		amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ
						DDESGQKKLHGLQAILVHEASGTTAITATAT GYQESHLSSAR
76	1426	A	1155	38	410	PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK PDCKEIWIFWWGDEPNLVVQYIMNCMLWK KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
77	1427	A	1162	526	350	T RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	A	1171	1	1293	MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT
						TSPRNWCIKMVCNPWFECVSMLVILLNCVTL GMYOPCDDMDCLSDRCKILOVFDDFIFIFFA
- jk						MEMVLKMVALGIFGKKCYLGDTWNRLDFFI VMAGMVEYSLDLONINLSAIRTVRVLRPLKA
			ł			INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF
	1					FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PP\YYQPEEDDEMPFICSLSGDNGIMGCHEIPP
						LKEQGRECCLSKDDVYDFGAERQDLNASGL CVNWNRYYNVCRTGSANPHKGAINFDNIGY
}		)				AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI
ļ						YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN FLSATHLGGLFPPWPLVEERKLKPKASQQCPI
						CHKVIMGAGKLPRHMRTHTGEKPYMCTICE
			1	ļ		VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS AIWQQAREVVRFNGLEDRVHVLPGPVETVEL
		1				PEQVDAIVSEWMGYGLLHESMLSSVLHARTK VVKDGGFFLPXSSELFM
82	1432	A	1187	12	716	DFVDAARNLPLESTKSPAEPSKSVPSLE\DPRA
						SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH
			1			GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD
	ļ		1			TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP
	ľ					SQESNVSLSGSSRSALFERDDHGKAEAPSPSF DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC
						WGRGHGCGQEALSTSHGYHLFCALLTGFLFA SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF
84	1434	A	1192	45	476	Q   LGDVGFWVERTPVHEAAQRGESLQLQQLIES
						GACVNQVTVDSITPLHAASLQGQARCVQLLL AAGAQVDARNIDGSTPLCECLRLGQHRVCEA
				-		LAVLRGQGQPSPVHSVPPARGLHXREFRMC* GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR
			1			SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT
06	1426	<del> </del>	1215	3	405	GRSPCPSLPGTTRTNSLL LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC
86	1436	A	1215	,	405	NLGFFLEGHAVLTCHAGSENSATWDFPLPSC
L		Ц			<u></u>	RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

NO: of much exide exide sequence where the control of the control		VE		1000	7	N - P 3 3	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide sequence with the sequence s	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
Section   Sect			hod	f .			E-Dh-vuldenine G-Gheine H-Histidine
1437   1216   226   964   96	-						I-I-almoine Vel voine I el guine
14							
### ### ### ### ### ### ### ### ### ##	-	uence		1			
	uence	1	ļ	914			
peptide		1	1	1			
		Į		ļ.		sequence	Y=1yrosine, X=Unknown, *=Stop codon,
NADCTWTILAELGDTIALVFIIPQLEEGYDFL		l .	ļ	1			
### 1437 A 1216 226 964 GTARPGMYGFGANRAGRILPSLVLGYLLV VTVLAFNYWSISSRIVLLGEVAELQGOVQ REPVARGRILERNSDLFAVWGHAGPTORPEG GRIPPPQCPAAGQRGPEEMBEDKVKLQNN ISYGMADDHILKEGLAERGEFIRQEGOLD YRKNNTYLVKRLBYESPOCOCQMKELRAGH BENIKKLADQHLEGKQETQKIGGNGKELDI NQVYPKNRFVARNVADKNEEPSSNHDHG PROFITS OF THE STANDARD REPORT			ļ		sequence		
1437							NADCTWTILAELGDTIALVFIDFQLEDGYDFL
			i				
	87	1437	A	1216	226	964	
RTEVARGRIEKRNSDLFAVVGHAQETIDREG   GILRPQQPAAGGGPREEMBIDEVKIQNN   ISYGMADIHILKEQLAELRQEFIRQEDQLQ    YERNITYLVKRIEVESPQCQGAQMERLRAQD    YERNITYLVKRIEVESPQCGAQMERLRAQD    YERNITYLVKRIEVESPQCGAQMERLRAQD    YERNITYLVKRIEVESPQCGAQMERLRAQD    YERNITYLVKRIEVESPQCGAQMERLRAQD    YERNITYLVKRIEVESPONHTPHO    YEVARAD   YEVAREVENITY   YEVARAD   YEVAREVESPNHTPHO    YEVARAD   YEVAREVESPNHTPHO    YEVARAD   Y	•			1	į	ļ	VIVVLAFNYWSISSRHVLLQEEVAELQGQVQ
		ļ	ļ	}	ļ	)	
		1	ì			ļ	GRLRPPOOPAAGORGPREEMEDDKVKLQNN
				ļ	ļ	1	
BENIKKLADOPLEGOKOĞTÖKIĞISDÜKELDI   NOQVYEKIPEY VARIN VADKISEPSINHEPIĞ    88   1438   A   1218   1   534   PEFGTTISCGYLMATDVSRR9SVHKAVELEĞE   RVKSAĞAWILIPYSDERFYWDLIKLMVGN   LIVLPYGITFIKEENSPPWIVFIKVLSDITFILD   LVINIRTGIVVEBĞABILLARABIRTIXLTW   FLVDÜLSSIPYDYPTILVVELEPERLDASVYKTAR   ALBIVRFTKILSLIRL   ROM			ł	1		ł.	
NNQVYENIPKY NABNYADKNEEPSSNIHHHG    88			1				
1438			,		ļ	j	
RVKSAGAWIIHPYSDFRFYWDLINLLIMVGR LIVLPYGITFFKEENSPHWVFNVLSDIFFILD LVLNFRTGIVVEEGABILLAPRAIRTRYLRTW FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR RARIVRFTKILSLLRL  89 1439 A 1223 1 743 MGFDEVFMINLRRQDRRERMLRALQAGEIE CLVVAVAPRFVDADNLILNPDTLSLLIAENTTVV APMLDSRAAYSIPWCGMTSNAAPGRILAAMLET LVVVAPRFVDADNLILNPDTLSLLIAENTTVV APMLDSRAAYSIPWCGMTSOQYYKRIPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL UAFYPPHPDYTWSFDDIIVFARSCKQAEVQMY VCNREEVGFLPVPLRAHSILQDEAESRHHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL UKTSIFFYLKNIVVADLIMTLITFPRIVHDAGF GPWDPKFILCRYTSVLFYANMDTSIVLGLITY UNDRYMKVYRILMDSWMTGIISFTRVYLLG LGARLVWFGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSSLGAESPSRMVASQAYNLT SALTPLITRSVLINEPHLAGSRAPANLSD VVQLIFLVDSNFFPFQ YISNYTVSTKVASMAF QTQAGAQUPERLASERATIVKVYPNSDWAAR GHRSSANSVVQPQAFVGAVVILDSSNPAAV LHCLOLNYTLLDGRYLLSEPPLAVVLHSEPR PNEHNCSASRRIPPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVCLYTSLC QYFSEEDVVWRTEGLLJEETSPRQAVCLITR HLTAFGTSLFVPPSHIRFVPPETADVNYIVML LQTARSTFFLVND WLSVETEANGGLVEKEVL AASKASFRVPTPSAALHKILDQLDASRGRAP FCGQRGRKYEILVKTGWGRGSGTAHVGM LLYGVDSRSGRHHD LOQDAFFIRNSILDFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVND WLSVETEANGGLVEKEVL AASKASFRVPTPSAALHRRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCVLLICL FLGANAVWYGAVCIDSAYSTGRVSCINPLSV DTVAVGLVSSVVVYPVYLAILFIFMSRSKV GWGWGPGSTONGA WSAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VPDEENILNELNDFLREEINVFNCKLVATMP LFANADPHVTVAMSKLKKFEVFQPGDYIIREG AVGKMYFTQHGVAGVITKSSKEMLTDGS YGGEICLLYGRGRTASVADTVCRLYSLSVD NPNEVLEEYPMMRRAFETVADBLIDRICKEN LFANADPHVTVAMSKLKKFEVFQPGDYIREG AVGKMYFTQHGVAGVITKSSKSKRLTDGS YGGEICLLYGRGRTASVADTVCRLYSLSVD NPNEVLEEYPMMRRAFETVADBLIDRICKEN SILLQKFQKDLNTOVFNNQENELLKQIVKKN SILLQKFQKDLNTOVFNNQENELLKQIVKKN KAEPPERKRKDSTSTFNNQFSALHTNYVYCT KAEPPERMSTRINTVYCT  93 1443 A 1249 180 901 TVPPPFGGFSFPALHEFKRSTSTFORAELKEERL PGRKASCSTAGSGRGLPPSSRWTRNTVYVCT	-00	1420	+	1010	<del> </del>	524	PRECITISCOVI MATDVSRPPSVHKAVEIROF
LIVLPYGITFYEEDSPPWIVPNVLSDIFFILD	88	1438	A	1210	1	334	
By   1439   A   1223   1   743   1245   1   743   1246   1246   1246   1442   A   1246   5   1249   180   901   1779FF9F9F1FILDREPLYMENT   1247   1248   1249   180   901   1779FF9F9F1FILDREPLEMENT/OR   1247   1248   1249   180   901   1779FF9F9F1FILDREPLEMENT/OR   1247   1248   1249   180   901   1779FF9F9F1FILDREPLEMENT/OR   1247   1248   1249   180   901   1779FF9F9F1FILDREPLEMENT/OR   1248   1249   180   901   1779FF9F9F1FILDREPLEMENT/OR   1248   1249   180   901   1779FF9F9F1FILDREDLEMENT/OR   1249   180   901   1779FF9F9F1FIRSRSALPFISTISTALERER   1248   1249   180   901   1779FF9F9F1FIRSRSALPFISTISTALERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERERER   1248   1246   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F9F1FIRSRSELERER   1248   1249   180   901   1779F9F9F9F9F1FIRSRESHERER   1248   1249   180   901   1779F9F9F9F9F1FIRSRESHERER   1248   1249   180   901   1779F9F9F9F9F9F1FIRSRESHERER   1248   1249   180   901   1779F9F9F9F9F9F9F9F9F9F1FIRSRESHERER   1248   1249   180   901   1779F9F9F9F9F9F9F9F9F9F9F9F9F9F9F9F9F9F9		į.	1	1		l l	
		)	}	}	1	1	
ALRIVAFTKILSLIRL		1	1	1	Î	1	
1439		1	1	1	1	1	
CELVEAVDGKVGMLTRSNAPGRHLAMLET LVVVAPRFVDADNJLINPDTLSLIJAENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PERKEDRRGCFAVPMYHSTPLDLEKAASRNL VAPYPHPDYTWSPDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISVI.  90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMILTFPFRIVHDAGF GFWDFKFILCRYTSVLFYANMDTSIVVLGLII/ YDRYWKVVRILLDSWMTGIJSFTRVYLLG LGARLWWFGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNFFPFGYISNYTVSTKVASMAF QTQAGAQPIERILAGRAFRATVKVFYNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PMEHNCSASRRIPSSLQGADHRPYTFFISFGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEDVVWRTEGLLPLETSFRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAALHKLDQLDASRGRAIP FCGQRGFKYEILVKTGWGRGGSTTAHVGIM LYGVDSRSGHRHLDGDRAFRNSLDIFQIATP HSLGSMWKRVWHDNKGLSPAWFLQHILVRG LYGVDSRSGHRHLDGDRAFRNSLDIFQIATP HSLGSMWKRVWHDNKGLSPAWFLQHILVRG LYGVDSRSGHRHLDGDRAFRNSLDIFQIATP HSLGSMWKRVHDNKGLSPAWFLQHILVRG FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYRVYLAILFRMSRSKV DTVAVGLVSSVVVYPVYLAILFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGYHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCKKLVATMP LFANADPNFVTAMLSKLRFEVPQFGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTVROPRSEMVSSAHDPN KAEPPERADSTSTPNNLPPSSMYMTRNTTYVCI  93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSGRGLPPSSPMVSSAHRPN		1	1				
LIVVVAPRFUDADNI.II.NPDITI.SLLIAENKTVV	89	1439	A	1223	1	743	
APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKTDRRQCFAVPMVHSTFLIDLRKAASRNL NAFYPPPPDYTWSTPLIDLRKAASRNL NAFYPPPPDYTWSTPLIDLRKAASRNL VAPYPPPDYTWSTDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEWMYPSSPSSAQSMAVVSADHIGLVISVI.  90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMITFPFRIVHDAGF GFWDFKFILCRYTSVLFYANMDTSIVVLGLII/ YDRYWKVVRELUDSWMTGUSFTRYYLLG LGARLVWFGKLILAKGGHGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPLITRSRVLNEPHTLAGPGSRAPANLSD VVQLIFLVDSNFPFPGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVYTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIPPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLETSFRQAVCLTR HTAFGTSLFVPPSTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCQQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGYDSRSGIFRHDODRAFHRNSDIPQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTFSAALLRFRRLDVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPQPGDYIREG AVGKKMYFQHOXQVITKSSKEMKLTDGS VFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETYAIDRLDRIGKKN SILLQKFQKDLNTGVFNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHEKRSPTSTGBAELKEELL PGRKASCSTAGSGSGRGLPPSSPMVSSAHNPN KAEPPERRADSTSTPNNLPPSMMTRRNTTYVCT		<b>\</b>	1			i	
PIRKADRRGCFAVPMVHSTELDILKAASRNIL	,	j.	-	1	-	1	LVVVAPRFVDADNLILNPDTLSLLIAENKTVV
1440   1227   2   349		1	1		Į.		APMLDSRAAYSNFWCGMTSQGYYKRTPAYI
90 1440 A 1227 2 349 NKTSFIFYLKNIVADLIMILTFYFRIVHIDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHLAWDSWMTGIJSFTRVYLLG GARLVWFGKILGAKGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGPISRAPANLSD VVQLIFLVDSNPFFFGYISNYTVSIKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEFFEPYLAVVLHSEPR PNEHINCSASRRIRPESLQGADHRPYTFFISFGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVYPPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCQQRGRKYKELVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWIDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTTSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPPSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRINPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGPGGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVMFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFQHGVAGVAGVASVASPCPPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVMFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENELKQIVKH  93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPPSSPMVSSAHNPN KAEPPERKDSTSTPNNLPPSSMMTRNATTYVCT			1	l	1		PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL
LEYMYPSSPSSAQSMAVVSADHIGLVISYL	}	1			}	1	\AFYPPHPDYTWSFDDIIVFAFSCKQ\AEVQMY
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMILTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG LGARLVWFGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILITSRVLNEEPLITLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKLYPINSDWAAA HLQLNYTLLDGRYLSEEFEPYLAVYLHSEPR PNEINCSASRRIPESLQGADHRPYTFFISPGT RDPVGSYRLINSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRTYPFEPYTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRWHDNKGLSPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTFSAALRFRILLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYYVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDFLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFOPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSFTSTGAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	l	1		ł		l	VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ
GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/STRVYLLG LGARL/WPGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRV/LNEEPLTLAGFSRAPANLSD VVQLIFL/DSNPFPFGYSINTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSSSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFTPLVDWLSVETEANGGLVEKEVL AASKASFTPTSVALLKFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYLAILFIFRMSRSKV GWGWGFGTGNGAWASAPCPEEPLSSAAAR GKGYHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPQFGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEPYMMRRAFETVADIRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN				Į.	1		LEVMVPSSPSSAQSMAVVSADHIGLVISYL
GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/STRVYLLG LGARL/WPGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRV/LNEEPLTLAGFSRAPANLSD VVQLIFL/DSNPFPFGYSINTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSSSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFTPLVDWLSVETEANGGLVEKEVL AASKASFTPTSVALLKFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYLAILFIFRMSRSKV GWGWGFGTGNGAWASAPCPEEPLSSAAAR GKGYHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPQFGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEPYMMRRAFETVADIRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN	90 .	1440	A	1227	2	349	NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF
91 1441 A 1245 3 1937 LGSSDVRAPORSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFOSRAPANLSD VVQLIFLVDSNIPFFGYISNYTVSTKVASMAF QTQAGAQPIERLASERATITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFK YEIL VXTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWRRVWHDNKGLSFAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FPDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDVIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN	100	1	1	-=-/		1	GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/
LGARLVWFGKLILAKGGHGGISWL		İ	1	ł.	1		
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRS.VLNEEPLILAGF\SRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSV\VQPQAFVGAVYLDSSNPAAV LHLQLNYTLLDGRYLSEEPEYLAVYLHSEPR PNEHNCSASRRIRPESLQADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIUYPD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FPKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRRLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH SILLQKFQKDLNTGVFNNQENEILKQIVKH PGRKASCSTAGSGSGRJPPISSPMVSSAHNPN KAEPERKDSTSTPNNLPPSMMTRRNTYVCT	j	j	1	j	j	1	
SALTPILTRSRVLNEEPLTLAGF\SRAPANI.SD VVQLIFLVDSNFFFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPINNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFYFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHHWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEEMILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVADDRJGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGFSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERKDSTSTPNNLPPSMMTRRNTYVCT	01	1441	<del>- </del> -	1245	<del> </del>	1027	
VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAP FCGQRGFKYELVXTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFFIRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVTPSVAALLRFRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDFLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERKDSTSTPNNLPPSMMTRRNTYVCT	1 91	1441	A	1243	3	1937	
QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSNAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGBSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQFGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH PGRKASCSTAGSGSRGLPPSSFMVSSAHNPN KABIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	ļ		1	1	
GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYVSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	İ		1	1	OTO A CACIDIEDI A SED AITVK VENNISOWA AR
LHLQLNYTLLDGRYLSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAPP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVODSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFIFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		i		1			CLIBGG VICTORO VENCY VICTI DESIBA VI
PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATT HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPNSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	1		}	ļ	}	
RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1	}		1	
QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVVLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	ĺ	1					
HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFK YEIL VKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGP FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	[		1	1	}		
TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	1	1 .	1	!	1 -
FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1 -	-			
LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGP FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1		1		
HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGP FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NPNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	1		i	1	FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM
LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	{				1		LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP
AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1.	}	1		
FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	1	1	1	1	LQTARSTFFLVNDWLSVETEANGGLVEKEVL
FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1	1	1		AASKASFRVPTPS\AALLRFRRLLVAELQRGF
FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1		1	1		1	FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL
DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	1	1	Ì			FLGANAVWYGAVGDSAYSTGRVSRLNPLSV
GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	ľ	1		1		DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							GWGWGPGSTGNGAWASAPCPEPPLSSAAAR
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFOPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NNFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	}	}		1		}	GKGVHORLLGKGOHT
LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH SILLQKFQKDLNTGVFNNQENEILKQIVKH PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	02	1//2	+ 🔭	12/6	- 5	562	VEDEENILNELNDPLREEIVNFNCRKT.VATMP
AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	74	1442	, A	1240	1		I FANADPNEVTAMILSKI REEVEOPGDYIIREG
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1		-	1	1	ł	
NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1			1	1	1	
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1					1	
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			I	1	1	1	NTNEVLEE I TMMKKATE I VAIDKLUKIGKKN
PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							SILLQKFQKDLNIGVFNNQENEILKQIVKH
KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	93	1443	A	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
			-	}	ì		PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN
	1		}		1	1	
			1	-	1	1	ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

MO: of mucle colide colide colide colide colide sequence   Mose of mucle colide colide sequence   Mose of mucle colide seque	000 ID	I CEO III	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence peptide colidide location (1946) and the provided sequence peptide sequ	SEQ ID	SEQ ID					
Section   Sect		1	liou l				
### 1446 A 1294 1 1456   1456   1467   147   147   A 1295   2 2057   1447   A 1295   2 2057   1447   A 1295   2 2057   1448   A 1304   118   453   1660   16			ł	1			
144				1			
amino acid or festion of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per se		401100					
residue of peptide sequence   Y=Tyronine, X=Unknown, **Sipp coton, / Possible muleotide delicion, / Possible muleotide delicion,   Possible muleotide delicion,   Possible muleotide delicion,   Possible muleotide delicion,   Possible muleotide delicion,   Possible muleotide delicion,   Possible muleotide delicion,   Possible muleotide   Possibl	duise	İ	1			of peptide	T=Threonine, V=Valine, W=Tryptophan,
		-	1		residue of		
		Ì	1		peptide	•	/-possible nucleotide deletion, /-possible
1444			1			ļ	nucleotide insertion
1444		<del> </del>	<del>                                     </del>	1			LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG
1444   A   1261   3   385		ì	ļ		ļ	j	GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP
1444   A   1261   3   385		'	İ			1	TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP
1445   A   1282   2   250   3   3   3   3   3   3   3   3   3		·		ĺ			RQEDHLSPGGRGCSEL
1445   A   1282   2   250   350	94	1444	A	1261	3	385	KFSQWGLTKPKLSNASP/WISLVKKLMKKWS
DPNHFYAMDETHHIKCLVLRIQEAFGNKLCPA				}			VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ
1445   A   1282   2   550   GPRDNPGNEDPRFEIVEHFGJAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPTLANLGRVAQVLRLMRIFRILLKARH STGLRSLGATLKYSYKEVGLLLTVLSVGISIPS VVAYTIEKEENBGLATIPACWWAATVSMTT VGYGDVVPGTTAGKLTASALTIPACWWAATVSMTT VGYGDVVPGTTAGKLTASALKGKFGNDFEFG VAYTIEKEENBGLATIPACWWAATVSMTT VGYGDVVPGTTAGKLTASALKGKFGNDFEFG VGYEKFFDANCNEKDYNPVAAGGGGETEVAP SIVAPUNKEPANPAGGSAKAGKFGNDFEFG VFKKFFDANCNEKDYNPVAAGGGGETEVAP SIVAPUNKEPACCEGGVCKVARPNNGYT SFDTFSWAFLSIFRLMTQDYWENLYQLTLRA AETTYMIPLVALVILLIGSLYLVTLLAVVAMA YEEQONQATLEAAGRENPNGYT SFDTFSWAFLSIFRLMTQDYWENLYQLTLRA AETTYMIPLVALVILLIGSLYLVTLLAVVAMA YEEQONQATLEAAGRENPRGSGEKDED EFGKSSEDSIRRKGFRSIEGRRLTYSKRYSS PHOSLISIRGSIFSPRNNSTSIFSFRGRAKDV GSADQAATATASEHSREPSAAGRLSDSSEAS KLSSKSAKERRNRKKKRKGSGGEKDEDD EFGKSSEDSIRRKGFRSIEGRRLTYSKRYSS PHOSLISIRGSIFSPRNNSTSIFSFRGRAKDV GSENDADDEHSTFEDNESRRDSLFVPRHGE RRNSNLSGTRSSEMALAVFRAYBAGE SEPGKSSEDSIRRKGFRSIEGRRLTYSKRYSS PHOSLISIRGSIFSPRNNSTSIFSFRGRAKDV GSENDADDEHSTFEDNESRRDSLFVPRHGE RRNSNLSGTRSSEMALAVFRAYBAGE SEPGKSSESDSTRKGFRSIEGRALTYSKRYSS PHOSLISIRRSSINGHTFTGRATED ON THE STANDARD SEPGKSSESTIFSTSFSVAMSANCHMINTOF CTSPVDIDDASSPPSGRIPSTTIFTGQGALS SEPSTSPSSVALSSRSVHKFFTGRFTEDD MET SEPDELSSPFSGSPSKAMSANCHMINTOF CTSPVDIDDASSPPSGRIPSTTIFTGQGALS SEPSTSPSSVALSSRSPSGRIPPLD MET SEPTEMBELSVLESCHWASTITTOR CTSPVDIDDASSPPSGRIPSTTIFTGQGALS SEPSTSPSSVALSSRSVHKFFTGRFTEDD DATA AVTRSSWELVRSFGGFKRASANCLMYNN DKAGFLIGEDILKTGQFLVGIETLPPDLRFV EEDNGRFFKELEEWDAQLAKALGEKLLAS QKLRSSTSVTTAQAAGDFKYLEOFSBFSK HILKETTGITTKASHEHDENSFTLUGAIKLE YARLVKLAQEDTPPETDVRLHHVVVYFIQNQ APKKILEKTLLEGFGDRNLTSGRCHIMMSVA QKALEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRSSTJOLLICHAYONKELL SKGLYRGHGDELISHYRECLIKLINGAAEL ESGEDRSVNNGLIMNSETVPFLILLLVDEME KDILLAVENGRWWCSLLLGFGGDRNTSEFTLLCHRENGAAER VARLAKLLELGFGDRNTSGRCHIMMSVA QKALEMIKPEVSTHELE SEGEDRSVNNGLIMNSETVFPLILLUPDEME KDILLAVENGRWWCSLLEGGERGENDARVA GKALEMIKPEVSTHELE SEGEDRSVNNGLIMNSETVFPLILLUPDEME KDILLAVENGRWWCSLLGGGRFFALGKLIT DFLYKLLGCROPPATALGKLIT DFLYKLLDCSMEKSFFEPRLGSSYRLD SIGNATHANGE VGRANGALGRV WSOLLLUVRDGWUQLLDETKEELDSYRLD SIGNA		1	[				EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL
1445   A   1282   2   550		1	i	i	Į.		DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA
VAPDFLKFFKNALNLIDLMSIVPFYITLVVIL   VVESTPILANI, GRVAQVLIRMFIRLIKLARH     STGLRSLGATLKYSYKEVGLLLLY, SVGSIFS     VAYTIEKEENEGLATIFACWWATVSMT     VGYGDVVPGTTAGKLTASACLLA     GLIPPSINENAGLLVGRCLCSAALRPVGDLIT     SGQVAVRNAPQAGSAKAGKGFQDNEFFIQ     YFKKFFDANCNEKDYNPVAAGGGGETEVAP     SIVAPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNEYNOCEGYICVGKAGRENPINGY     STAPPUNKEYNEYNOCH     SELENGAR     S		1	1	1		1	CR
VAPDELKFFKNALNLIDLMSNYPYTILVVIL   VVESTPITANIGRVAQVLRIMFIFRILKLARH   STGLRSLGATILKYSYKEVGLLLI VLSVGISFS     VVAYTIEKEPENGLATIPAGWWAATVSMTT     VGYGDVVPGTTAGKLTASACILA     1456   QILPPSNRENAGLLVGRCLCSAALRPVGDLIT     SGQVAVRHAQOGSAKAGKFQDNEFEIQ     YFKKFFDANCNEKDYNPVAAGQGGETEVAP     SIVAPUNKRYNQCEGVICVCKAGRNPNYGYT     SFDTFSWAFLSLFRLMTQDYWENLYQULTRA     AETTYMIFALYVILLIGSLIVILILAVVAMA     AETTYMIFALYVILLIGSLIVILILAVVAMA     YEEQNQAT LEBABQKASFQQMLEQLKKQQ     EAAQQAATATASEHSREPSAAGRLSDSSSEAS     KLSSSASKERRNRKKKRKQEQSGEEKDED     EFQKSSSEDSIRKGFRSIEGNRLTYSKRYSS     PHOSILSIRGSLFSPRRNSSTSLFSFRGRAKDV     GSENDFADDEHSTFEDNESRRDSLFVPRHGE     RRNSLSQTRSSSMLAVFRNIGKMHSTVDC     RGVSLLVGFSVFTSPVQOLLPEVIIDKATD     DNGTTETEDMESRRDSLFVPRHGE     RRNSLSQTRSSSMLAVFRNIGKMHSTVDC     RGVSLLVGFSVFTSPVQOLLPEVIIDKATD     DNGTTETEDMESRRDSLFVPRHGE     RRNSLSGTRSSSMLAVFRNIGKHHSTVDC     RGVSLLVGFSVFTSPVQOLLPEVIIDKATD     DNGTTETEDMESRRDSLFVPRHGE     RRNSLSGTRSSSMLAVFRNIGKHHSTVDC     RGVSLLVGFSVFTSPVQOLLPEVIIDKATD     DNGTTETEDMESRRSSSHVSMMFTFORMILEDPSQR     QRAMSIASILTNITVE     GRAMSIASILTNITVE     GRAMSIASILTNITVE     SGLPSTSPSSVALSSRSVHKFTGSRIPPLD     MEPAPRHITEELSVLESCLIKWRTEIENDTR     DLQESISRHRTIELMYSDKSMQVPYRLBAV     LVHEQQANAGHYWAYIPIPTSTIEQQGALS     SELPSTSPSSVALSSRSVHKFFTGSRIPPLD     MEPAPRHITEELSVLESCLIKWRTEIENDTR     DLQESISRHRTIELMYSDKSMQVPYRLBAV     LVHEQQANAGHYWAYIPIPTLDLRFV     EEDNQRFEKELEBWAQLAQKALQEKLLAS     QKLRESSTSVTTAQAAGDPKYLEOPSRSPSK     HLKESTIQITIKASHEHDENSFWMKYNDI     AVTKSSWEELVRDSFGGTRNASAYCLMYIN     DKAQFLIQEDLIKTGQPLVGHITLPPLIRPF     EEDNQRFEKELEBWAQLAQKALQEKLLAS     QKLRESSTSVTTAQAAGDPKYLEOPSRSPSK     HLKESTIQITIKASHEHDENSFULQSAILT     YARLVKLAQEDTPPETDVRLHHLVVYTIQNQ     APKKILEKTLLEQFGDRNLSFDERCHNIMEVA     QKLRESSTSVTTAQAAGDPKYLEOPSRSPSK     HLKESTIQITIKASHEHDENSFULQSAILT     MTLICILENFQRSVTIDSLFLICKHNQAAEIL     SGGDRVNNGLIINNEFIVPTLLLLUVBME     KDILAVEDMRRWCSYLQGMPHLQEKLT     DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER     FARIMLSLSRTPADGR     SGCDRFTADGR     SGCDRFTADGR     SGCDRFTADGR     SGCDRFTADGR     SGCDRFT	95	1445	A	1282	2	550	GPRDNPG\EDPRFEIVEHFGIAWFTFELVARFA
STGLRSLGATLKYSYKEVGLLLLYLSVGISIPS VVAYTIEKERNEGLATREWWWATYSMTT VGYGDVPGTTAGKLTASACILA  96 1446 A 1294 1 1456 QLIPPSNRENAGLIVGRCLCSAALRPVGDLIT SGQVAVRAPQAGSAKAGKGKPQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGGETEVAP SIVAPVLNKPNQCPEGYJCVKAGRYPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AEITYMIFLVLVLUGSLYLVTLILAVVVAMA YEEQNQATLEBAGQKEAFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSKSAKARKRNRKKKKQKQSGGEEKDED EFQKSESEDSIRRKGFRSIEGNRLTYEKRYSS PHQSLLSBRGSLFSPRNSLSFSFRGAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RNNSLSQTRSSELSFSFRGAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSLSQTRSSELSFSFRGAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSLSQTRSSELSFSFRGAKDV GRAMSLASILTNIVE  97 1447 A 1295 2 2057 QTQLPTKSSQQLRKGGNCVRCKMQMFIAE EVLLKYRITFYNNNKGPNMLYHEIKAFVHFML NRYLSYGSGFKRFFLVDVLQYALEFASSKPV CTSPVDDDASSPSGSIPSQLFSTTEQQGALS SELPSTSPSSVALSRSSVHKFFTGSRIPPDLP MEPAPRHTTEELISVLESCLHRWRTEENDTR DLQESISRHRTELMYSMGQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI ANTKSSWEELLVRDSFGGYRNASACLMYIN DKAQFLIQEDLIKTGGPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALGEKLLAS QKLRSISFTSVTTAQAAGDFKYLEOFSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDITPETDYRLLOFKSRBF KHKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPETPTLLHVVYFIQNQ AFKKIEKKTLLEQFGDRNLSFDERCHNINKVA QAKLEMIRPEEVNLEEYEEWHQDYKRFEETT MYLIGIGLBRYRRECLIKLKRQAAEL ESGEDREVNIGLIMMPETVPFLPLLLVDEMEE KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPRLFSYSTHELCER KARIMLSLSKTPADGR SGGROFFONGLIMMPETVPFLPLLLVDEMEE KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPRLFSYSTHELCER FARIMLSLSKTPADGR SGGROFFONGLIMMPETVPFLLLHOPTUSH LMTCKQGGRQVQOFBOLAGKLFEMDAHGRV WSQDLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTICSYNSILS CGGYFCHTTCAPQAPPCVPPDLLKRLEGENTRA			1			1	
STGLRSLGATLKYSYKEVGLLLLYLSVGISIPS VVAYTIEKERNEGLATREWWWATYSMTT VGYGDVPGTTAGKLTASACILA  96 1446 A 1294 1 1456 QLIPPSNRENAGLIVGRCLCSAALRPVGDLIT SGQVAVRAPQAGSAKAGKGKPQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGGETEVAP SIVAPVLNKPNQCPEGYJCVKAGRYPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AEITYMIFLVLVLUGSLYLVTLILAVVVAMA YEEQNQATLEBAGQKEAFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSKSAKARKRNRKKKKQKQSGGEEKDED EFQKSESEDSIRRKGFRSIEGNRLTYEKRYSS PHQSLLSBRGSLFSPRNSLSFSFRGAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RNNSLSQTRSSELSFSFRGAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSLSQTRSSELSFSFRGAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSLSQTRSSELSFSFRGAKDV GRAMSLASILTNIVE  97 1447 A 1295 2 2057 QTQLPTKSSQQLRKGGNCVRCKMQMFIAE EVLLKYRITFYNNNKGPNMLYHEIKAFVHFMI NRYLSYGSGFKRFFLVDVLQYALEFASSKPV CTSPVDDDASSPSGSIPSQLFSTTEQQGALS SELPSTSPSSVALSRSSVHKFFTGSRIPPDLP MEPAPRHTTEELISVLESCLHRWRTEENDTR DLQESISRHRTEILMYSMGQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI ANTKSSWEELLVRDSFGGYRNASACLMYIN DKAQFLIQEDLIKTGGPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALGEKLLAS QKLRSISFTSVTTAQAAGDFKYLEOFSRSDFSK HLKEETIQITKASHEIHEDKSPETVLQSAIKLE YARLVKLAQEDTPFETDYRLHDVVVYFIQNQ AFKKIEKKTLLEQFGDRNLSPDERCHNINKVA QAKLEMIRPEEVNLEEYEEWHQDYKRFRETT MYLIGLGINFQRSSYUDSLLFILCAYQNNKELL SGGLYRGHDEELISHYRRECLIKLNEQAAGLE ESGEDREVNINGLIMMPSTVPFLPLLLVDEMEE KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER FARIMLSLSKTPADGR KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER FARIMLSLSKTPADGR KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER FARIMLSLSKTPADGR SGGROFFONGLIMMPSTVPFLPLLLVDEMME KOILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER FARIMLSLSKTPADGR SGGROFFONGLIMMPSTVPFLPLLLVDEMME KOILAVEDMRNRWCSYLGQEMEPHLGEKLT DFLPKLLDCSMEIKSFHEEPRLFSYSTHELCER FARIMLSLSKTPADGR SGGROFFONGLIMMPSTVPFLPLLLVDEMME LMTCKQGGRQQOPEDALQKLFEMDAHGRV WSQDLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS CGGYFCHTTCAPQAPPCVPPDLLKRLEGER FARIMSLSCOORSVQOPEDALQKLFEMDAHGRV WSQDLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS	(	1	1		1		VVESTPTLANLGRVAQVLRLMRIFRILKLARH
1446   A   1294   1   1456	ł	}	ł	1	1		
96 1446 A 1294 1 1456 QILIPPSINEENAGLLVGRCLCSAALRPVGDLTT SSGQVAVRNAPQAGSAKAGKGKFQDNPEFIQ YFKKFFDANCNIKRDYNPVAAGQGGETEVAP SIVAPVLINKPNQCPEGYICVKAGRNPHYGYT SPOTIFSWAPILAVIVILLGSLYLVTLILAVIVAMA AETTYMIRLVJVLILGSLYLVTLILAVIVAMA AETTYMIRLVJVLILGSLYLVTLILAVIVAMA YEEQNQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRRKKRKQKGSGGEKDEDD EFQKSESEDSIRRKGFFSIEGINRLITVEKRYSS PHQSLLSIRGSLFSPRRNSRISLFSPRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RINSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGGLLPVIIDKPATD DNGTTTETEMRKRSSFFIVSMOFILEDPSQ QRAMSIASILITIVE  97 1447 A 1295 2 2057 QPQLRKGGOVPCKMQMNFIAE EVLLKYRITFYNNINGPINNLYPEIKAFVHMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDDDASSPPSGSIPSGTLFSTTEQGGALS SELPSTSPSSVAJSSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIPDHRSRWMKYNDI AVKKSWELVRDSFGGYRNASAYCLMYIN DKAGFLIGEDLIKTGGPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDRKYLEQFSRSDFSK HILKEETIQIITKASHEHEDKSFETVLQSAIKLE YARLVKLAGEDTIPPETDYRLHHVVVYRIONQ APKKIIEKTLLEGFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLKLINEQAAGT ESGEDREVNNGLIINMFIVPFLPLLLVDEMEE KDILAVEDMRRWCSYLGGEMPHLGELT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSISRTPADGR  98 1448 A 1304 118 453 GGPSSRAIYLHRKEYSQNITSEPTLLQHRVEM LMTCKGGSGVYOGPDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNITCSYNSILS 199 1449 A 1306 3 1660 CGYPCHTTCAPQAPPCPVPPDLLRTALGVHPE	1		ļ			ļ	VVAYTIEKEEN\EGLATIPACWWWATVSMTT
SSGQVAVRNAPQAGSAKAGKGKPONNEFU YFKKFFDANCNEKDYPVAAGGGGETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA ABTTYMIPL/VI/VLIJGSLJVTLLLA/VYAMA YEEQNQATLEEABQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKKKKKGKGQSGEKDED EFQKSESEDSIRKGFRFSIEGRILTYEKRYSS PHOSLISIGGSLEFSPRINSITSLFSTRGRAEDV GSENDFADDEHSTFEDNESRRASDLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC ONGVSLVGGFSVYTSPVQQLLPEVIIDKPATD DNGTTTETEMKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 QTQLPYKSSQQLKKGGNCVRCKMQMNFIAE EVLKYRITFYNNNKGPNMLYEIKAFVHFMI NRYLSYGSGPKRFPLYDVLQYALEFASKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS. SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHTIEEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMMQVPYRLHAV LYHEGQANAGHYWAYPDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTQQPLVGIETJPPDLRDFV EEDNQRFEKELEEWDAQJLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETJQHTKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDITPETDYRLHHVVYYFQNQ APKKIEEKTLEQFGDRNLSFDERCHNIMKVA QKLREGETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETJQHTKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDITPETDYRLHHVVYYFQNQ APKKIEEKTLEQFGDRNLSFDERCHNIMKVA QKLLEGMKPEPVNLEEYEEWHQDVYRKPETT MYLIGLENFQRESVIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLIKLNEQAAELI ESGEDREVNNGLIMNREVNPFLPLLLVDEMSE KDILAVEDMRNRWCSYLQGEMEPHLQEKLT DFIFKLLDCSMELKSFEPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEN LMTCKGGSGVYQGPDALQKLFEMDAHGRW WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNITCSYNSILS		1	1	1		Į.	
SSGQVAVRNAPQAGSAKAGKGKFONNEFIL YRKKFDANCNREWPVAAGQGGETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA ABTTYMIPL/VLVILLIGSLYLVTLLLAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRRKKKKKGKEQSGGEKDED EFQKSESEDSIRRKGFRFSEGNRLTYEKRYSS PHQSLLSIRGSLFSPRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRASIFLFSRGRAKDV GSENDFADDEHSTFEDNESRRASIFLEDPSQR RRMSNLSQTSRSSRMLAVFPANGKMHSTVOL NGVVSLVGGFSVYTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITNITV QRAMSIASLITNITV DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITNITV HILLIAM SELFSTSPSVAAISSRSVHKPFTQSRIPFDLP MHAPRHTTEELESVLESCLHRWRTEIENDTR DLQESISRHRTTELMYSDKSMMQVYYRLHAV LYHEGQANAGHYWAYFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTQQPLVGIETLPFDLRDFY EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEFTQUHKASHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QKLEMKFPEVNLEFYEEWHQDVTKFRETT MYLIGLENFQRESVIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELI ESGEDREVNNGLIINNEFIVPFLPLLLVDEMSE KDILAVEDMRNRWCSYLQGEMEPHLQEKLT DFLPKLLDCSMELKSFEPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNITCSYNSLS  159 1449 A 1306 3 1660 CGYFCHTTCAPQAPFCPVPPDLLRTALGVHPE	96	1446	A	1294	1	1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT
SIVAPVLNIKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA ABTTYMIFL/IV/LILGSLYLVTLLLAV/VAMA YEEQNQATLEBAEQKEAFEQMLEQLKKQQ EAAQQATATAFSHEREPSAAGRISDSSEAS KLSSKSAKERNRKKKRKQKEQSGGEKDED EPQKSESEDSIRRKGFRPSIEGNRLTYEKRYSS PHOSILSIRGSLFSPRINSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFYPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVYTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMMFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHMI NRYLSYGSGPKRFPLVDVLQY ALEFASSKPV CTSPVDDIDASSPSGSPSSGTLPSTTEQQGALS SELPSTSPSVAAISSRSVHKPFTGSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIOVPYKHAV LVHEQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRTEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEDKSPETVLQSAKLEL YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEFYEEWHQDYRKFRETT MYLIGLENPQRSSYIDSLLFLCAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPPLPLLLVDEMBE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLDCSNBEKFHEPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRATVLHRKEYSONLTSEPTLLQHRVEH LMTCXQGSQRVQOPEDALQKL,FEMDAHGRV WSQDLLQVRDGWUQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS	**			1	1	1	SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ
SFDTFSWAFLSLFRLMTODVWENLYQLTLRA AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQOQATLEEAEQKEAEFQQMLEQLKXQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKAKERNRKKKRKQKEOSGGEKDEDD EPQKSSESEDSIRRKGFRFSIEGRRLTYEKRYSS PHQSLLSIRGSLFSPRRNSTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFYPRRHOE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVYSLVGGPSVTTSPVGQLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHPMI NRYLSYGSGPKRFELVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISRSSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIEADTR DLQESISRHRTEELSVLESCLHRWRTEIEADTR DLQESISRHRTEELSVLESCLHRWRTEIEADTR DLQESISRHRTEELSVLESCLHRWRTEIEADTR DLQESISRHRTEELSVLESCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLQESISRHRTEELSVLASCLHRWRTEIEADTR DLAGSGPKRASARYCLMYIN DKAQFLQEDLIKTGQPLVGIETIPPDLRPDV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESSTISVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVYYFIQNQ APKKIEKTLLEGFGDRNLSFDERCHIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKPRETT MYLIIGLENPQRSSYIDSLLICAYQNNKELL SKGLYRGHDEELSHYRRECLLKLNEQAALEH ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLQGEMEPHLQEKLT DFLFKLLDCSNEIKSFHEPPKLPSYSTHELCER FARIMLSLSKTPADGR  98 1448 A 1304 118 453 SGPSSRAIVLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQOPEDALQKLFEMDAHGRW WSQDLLQVRDGWLQLDIETKEELDSYRLD SQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		1		}	1	}	YFKKFFDANCNEKDYNPVAAGQGQETEVAP
AETTYMIELVAVILLAGI. YUVTLLAV/YAMA YEEQNQATLEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERNRRKKKKÇKEQSGGEKDED EPQKSSEDSIRKGPRISIGRRI. TVEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTTEDNESSRDSL. FVPRRHGE RRINSIL, SYSTESSEMLAVFPANGKMHSTYDC NGVVSLVGGPSVPTSPVQQLLPEVIDKPATD DNGTTTETEMKKRSSFHVSMDFLEDPSQR QRAMSIASILTNIVE  1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVILKYRITFYNNKGPML YIELKAFVHFM NRYLSYGSGFKFPLVDVLQY ALEFASSKPV CTSPVDDIDASSPPSGDFSQTLPSTIEQQGALS SELPSTSPSSVALSSRSVHKFFTGSRIPPLLP MHPAPRHITEELSVLESCLHRWRTEIEADTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLQEDLIKTGQPLVGIETIPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESSTSVTTAQAAGDFKYLEQPSRSDFSK HLKESTQUTIKASHEHDDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRINSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENPQRSSYIDSLLTLACAYONNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAEIL ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGGEMEPHLQEKLT DFLPKLDCSNEKFIEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRANTHREEYSONLTSEFTLLQHRVEH LMTCKQGSQRVQOPEDALQKLIFEMDAHGRW WSQDLLQVRDGWLQLLDIETKEELDSYRLD SQAMNVALNTCSYNSILS  199 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			İ	1 :			SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT
YEEQNQATLEABQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSEAS KLSSKSAKERNRRKKKKÇKQSGGEEKDED EFÇKSESEDSIRRKGFRFSIEGTRLTYEKRYSS PHQSLLSIRGSLFSPRRNSTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFYPRRIGE RRNSNLSQTSRSSRMLAVPPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPVIUKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSLASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLEKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVALSSRSVHKPFTQSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDERESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQENDLKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTAQAAGDPKYLGPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVALAQEDTPPETDYRLHHIVVYYFIQNQ APKKIEKTILEQFGDRNLSTBCRCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFILCAYQNNKELL SKGLYRGHDEELISHYRRECLLKINEQAAELF ESGEDREVNNGLIIMMEFIVPFLPLLVDEMEE KDILAVEDMRNRWCSYLQQEMEPHLQEKLT DFLPRLLDCSMEKSFHEPPKLEKLFNSTSTELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQQSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 199 1449 A 1306 3 1660 CGYFCHTICAPQAPPCPVPPDLLRTALGVHPE		[		ĺ			SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA
EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKKKQKEQSGGEKDED EFQKSESEDSIRKGFRFSTEGORLTYEKRYSS PHQSLLSIRGSLFSPRRNSRISLFSFRGAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVQQLLPEVIDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSLASILINIVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRPFLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHTIEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEEL VRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEBWDAQLAQKALQEKLLAS QKLRSETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVYYFIQNQ APKKIEKTILLEQFGDRNLSTDERCHNIMKVA QKLEMIKPFEVNLEFYEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKEIL SKGLYRGHDEELISHYRRECLLKINEQAAELI ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAYLHRKEYSQNLTSEPTILQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTSYNSILS 99 1449 A 1306 3 1660 CGYFCHTCAPQAFPCPVPPDLLRTALGVHPE			1			1	AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA
KLSSKSAKERNNRKKRKOKEQSGGEEKDED EFQKSESEDSIRRKGFRPSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKVRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MIPPAPHITEELSVLESCLFRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAVIPDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSYTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNA APKKIIEKTILEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVINEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELE ESGEDREVNNGLIMNEPIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DPLPKLLDCSMEIKSFHEPFKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLIFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  199 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	}		1	1		1	YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ
EFQKSESEDSIRRKGFRSIECNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRISLFSFRGRAKDV GSENDFADDEHSTREDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVILKYRITFYNNNKGPMMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSSIPSQTLPSTIEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHTEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKILAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQDTPPETDYRLHHVVYYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYYLIGILENFRESSYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMSETVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMPPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQQPEDALQKLJEEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		ì	1			ļ	EAAQQAATATASEHSREPSAAGRLSDSSSEAS
PHÖSLLSIRGSLFSPRNSRTSLFSFRGRAKDV GSENDFADDEHSTREDNESRRDSLFVPRRHÜG RRNSNLSQTSTRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTIVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYNI DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEOPSRSDFSK HIKEETIQITKASHEHEDKSKPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGILENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRCCLKLNEQAAELI SKGLYRGHDEELISHYRRCCLKLNEQAAELE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SQPSSRAIVLHRKEYSQNLTSEPTILLQHRVEH LMTCKQGSQRVQQPEDALQKLFEMDAHGRV WSQDLILQVRDGWVQLLDDIETKEELDSYRLD SQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE				İ			KLSSKSAKERRNRRKKRKQKEQSGGEEKDED
GSENDFADDEHSTFEDNESKRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSFVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSISILTNTVE  97 1447 A 1295 2 2057 QTQLPTKSQQLKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGGPKKRPLVDVLQYALEFASSKFV CTSPVDDDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEIDLIKTGQPLVGIETLPPDLRDFV EEDNQRFKKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQYBRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDITPETDVRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGGLBNFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTILLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	l	Ì	1	1	1	Ī	EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS
RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGYVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTTTETEMRKRSSSFVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 QTQLPTKSSQQLKKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAVCLMYIN DKAQFLIQEUDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKILAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMRPEGVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLOHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMVALNTCSYNSULS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	Ì	ļ		Ì			PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV
NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATID DNGTTTETEMRKRSSFHVSMDFLEDPSQR QRAMSLASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLVIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEUDIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKITLEQFGORNLSTDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFILCAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFI_PKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIM_SLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQYRGWUQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1	1	1				GSENDFADDEHSTFEDNESRRDSLFVPRRHGE
DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILITIVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLIKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNEELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWUQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1		1	Ì			RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRPPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVALISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYNI DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAILLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQOPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1		ł	(	Į.		NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVILKYRITFYNNIKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTIFQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMK YNDI AVTKSSWEEL VRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRSSYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR FARIMLSLSRTPADGR FARIMLSLSRTPADGR LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS	· ·	1	1	Ì	}	1	-
EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTIEQQGALS SELPSTSPSSVAAISSRSVHKFPTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKESTIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	İ	)	]	}			
NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRTPDLP MHPAPRHITEEEL\$VLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQOPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	97	1447	A	1295	2	2057	
CTSPVDDIDASSPPSGIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVITAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVYYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKET DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1				
SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYNN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDITPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLLCAYQNNKELL SKGLXYEGHDEELISHYRRECLLKLNEQAAELL ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		1.	ļ	,	1	Į.	
MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCVPPDLLRTALGVHPE			j		1	)	
DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTILLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTFADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1	- [			SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP
LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1		1	1	MHPAPRHITEEELSVLESCLHRWRTEIENDTR
AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1	l	1		
DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HIKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEPIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	)				1	}	
EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1		1		
QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVYYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		1					
HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		ł			{	1	
YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1	}			1	l	
APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQOPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1	j				1	
QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQQPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1		1	1	
MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1	1	-				APKKIEKTLLEQFGDRNLSFDERCHNIMKVA
SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1		1	1	ì		
ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	}	]	1				
KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		1	1		1	1	
DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1	1	1	{	{		
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	}		1	- {	1	1	
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					1		
LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1		1		<b>.</b>		
LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	98	1448	A	1304	118	453	SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH
SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					1		LMTCKQGSQRVQGPEDALQKLFEMDAHGRV
99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1			1	j	]	WSQDLILQVRDGWLQLLDIETKEELDSYRLD
	1	<u> </u>			1		SIQAMNVALNTCSYNSILS
	99	1449	A	1306	3	1660	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE
	1		[				TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

			Lone	<b>T</b>	Design 1 and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN 09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	·		314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	<u>'</u>	,		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ĺ	}	peptide	}	/-possible nucleotide deletion, \-possible
			]	sequence		nucleotide insertion
		ļ	<del> </del>	- Coquestor	<u> </u>	DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF
		]	]		ļ	SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP
		ļ				TTCTVLLLAESEGERERWLQVLGELQRLLLD
	1			ŀ	1	ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD
		1		ł	1	QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ
	}	ł	1	1	ł	QLTLSPSAGLLVVLCGRGPSVRLFALAELENI
					1	EV\EVPKIPESRGCQVLAAGSILQARTPVLCVA
	1	1	1			VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ
	)	}	1	]	1	SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL
		1	1	ł		GAGLVPEELPPSRGGLGEALGAVELSLSEFLL
		1	1	ŀ	1	LFTTAGIYVDGAGRKSRGHELLWPAAPMGW
				1		GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE
		İ		1	•	KNVRPENPEGSEFE TGTERVALTTERIQUAL KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE
		}			j	EQQKQQRREMLKDPFVRSKLISPPTNFNHLV
	1	1	ı			HVGPANGRPGARDKSP
	1	<del>  </del>	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA
100	1450	A	1318	910	190	LNDSMINETARDAARVQVASTLSVLVGLFQV
	1	1	1		1 .	GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF
		1	1	1	1	VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL
		ļ				POSKVGTVVTAAVAGVVLVVVKLLNDKLQQ
		1		1	1	QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\
						PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK
		1	1	1	1	IFALRHGYRVDSNQVWVMRDV
101	1451	A	1353	220	445	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR
		1	<b> </b>	<b>\</b>		SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD
	1	1	}			HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN
	}				1	WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY
		1	1	1		CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA
		1		{	ł	VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW YNLILTAULCLPLVIVTLCYTTIIHTLTHGHAN
				<u> </u>		\DSCLKQKARRLTILLL CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C
103	1453	A	1371	2	410	SFNEHGYHLFQAMRLAVEEINNSTALLPNITL
		1	ĺ	ļ	}	GYOLYDVCSDSANVYATLRVLSLPGQHHIEL
	ļ	1	ľ	1	1	QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP
	1	1		<b>\</b>	,	FLVPMLLEQ
104	1454	+	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG
104	1454	A	13/6	"	132	FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP
		1			1	RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW
	ļ	1	1	1	1	ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG
			1			KSREERFCNENTPCPVPIF
105	1455	A	1379	12	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD
103	1400	1 "	~~,	1-		IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF
	1	1			1	SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK
ļ	}	1			1	CQSDWMERWVDDAFWSFLF\SLILIVIMFLW
		1	1	1	1	RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW
	1	""			ŀ	QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG
		-				CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL
	1		1 .	1	1	FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF
						QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN
1	1.					YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM
	1	- 1			1	VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK
1						
	1	1	1	i		AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

SEQ ID	SEO ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciicc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	1		7.7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		]	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			}	peptide	004	/=possible nucleotide deletion, \=possible
	1	ŀ	Į.	sequence		nucleotide insertion
	<u> </u>		<del> </del>	soquence	<del> </del>	FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG
	Ļ			<b>\</b>	ì	FTYNVGALGGALAPIIGALIAQRLDLGTALAS
	l .	ļ	ì	!	į.	LSFSLTFVVILRNRRPGKSLVR
			<del>                                     </del>	10	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS
109	1459	A	1402	15	307	WSHNSNSMCWGKDQCPYSGCKEALIRTDGM
	ŀ		1	}	Ì	RVTSRKSAKYRLQGTIPRGDVSLTILNPSESDS
	1	l	1	1	1	GVYCCRIEVPGWFNDVKINVRLNLQRASTT
		<u> </u>				HEDLSSLLTRGSGNQERERQLKKLISLRDWM
110	1460	A	1421	3	350	LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF
	1			į	1	LAELAPPYGYLAICA SLLSC ICVILITOSCI
	1	ì	ļ	ì	1	FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL
	1	1	Ì			CLLLASSPFPLFILLASL
111	1461	A	1426	2	344	FTSTMTKPFEKESEQPA*ATLAFGAQTSTTAD
		1	1 .	1		QCALKPDLSYLNNSSSSSSTPATSAGGGIFGSS
	i	1	1			TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST
		1	ļ		1	SDSLLFSQDSKLATTS
112	1462	A	1434	46	372	TTSWTTSCTRSCT*SGASSGPGWTPRTTWWR
112	1402	1	1	1	1	SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC
		1		l .	1	STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC
			}	ì		MSSSTTSSTTSTF
	1463	A	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG
113	1403	1 ^	1437	13	1-7-	MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR
	1	1	ì	i		GIFLTFLLSNFLIVCVLLFYVSFYLFQSCINFVL
	1	<del> </del>	1,462	1	396	KQQAVPEPHSSTTTPQEQEQNWYGQDLLNLQ
114	1464	A	1463	1 1	370	QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP
	1		j		}	ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS
1		.   '	j			EDP*KNA*LKQMHAATTHWQQHQQHQVGC
	1		ļ		1	OYHGIMQ
						AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG
115	1465	A	1464	291	2	GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN
	İ	1	1	ì		MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN
	}	1	- }	1	1	
1	1	l _	1	1		NYCN
116	1466	A	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ
	1	1	- {	1	-	YWTKYQVWEWLQHFLDTNQLDANCIPFQEF
1	ł	1	1	1		DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ
1		- [	1			HLKWNGDSLFLCLSLPC
117	1467	A	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR
1		1				VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK
_		1	-	Į	]	QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN
1		1	1	(	ſ	NPEEELASDPNNEESL*RPWALEDFEIGRPLG
1		l l		ì	-	KGK
118	1468	A	1485	3	385	TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS
110	1400	1 ^	1405	1	1	PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL
1	1	1	ł	- [		OHPWIEGHTCLDNNIHQAASEPINNNFAESKR
1		1	1	l		NLAFLATGVVRHMRKLFMGANLEGPGPTVS
	1	-	Ì	i	(	н
112	-		1406	<del>                                     </del>	398	GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL
119	1469	A	1486	1	370	NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE
1		1				KALTKELKWVNWDLPOEAKOALELLGKWK
1		1		Į.	1	PMDVKDSLELLSSHYTNPTVRRYAVARLRQA
İ	1			1		DDEDLLMYL
						MGESPAV*GYFVLAGMNSAGLSFGGGAGKY
120	1470	A	1497	3	999	LAEWMVHGYPSENVWELDLKRFGALQSSRT
			1	}		ELAEWMYRU I FOENY WELDLARFOALQOOKI
	1	1	- [	1		FLRHRVMEVMPLMYDLKVPHWDFQTGRQL
		1	[	1		RTSPLYDRLDAQGARWMEKHGFERPKYFVP
ł	1	ł		- (		PDKDLLALEQSKTFYKPDWFDIVESEVKCCK
1	1		i	1 .	{	EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS
1	1	- 1				NDLDVPVGHIVHTGMLNEGGGYENDCSIARI
		1	1	ı	1	LAWDONER GOTTO ONGOWANT KKIMPKOSN
	}	}		1	1	NKRSFFMISPTDQQVHCWAWLKKHMPKDSN LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

					<del>5</del> 20 13 13	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	]		ļ	peptide	) Soquesion	/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
	ļ <u>.</u>	ļ	<del> </del>	Soquesion		MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
	i		ŀ			GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN
121	14/1	1 **	1	-		WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE
	ł	Ì	1	Ì	· ·	MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD
	İ	1	1			SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT
	}		{			WDPGTDTALGWSKQPSQSYTLFES*VGSGYII
	İ	ļ	İ			DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
		ł	1	1		RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA
	1	1	1	1		AKHGHSPAVQVLLAQWQDINEMNEKQQTPL
		}				HVAADRG
124	1474	A	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL
	ł	ì	1		j	HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP YVKFRLGHQKYKSKIMPKTLNPQWREQFDF
	(					HLYEERGGVIDITAWDKDAGKRDDFIGRCQV
		1		1		DLSALSREQTHKLELQLEEGEGHLVLLVTLT
	1 :	}		<b> </b>		DESALSREQIAKLELQLEEGEGGGCVCLVILI
	'	Į.				ASATVSISDLSVNSLEDQKEREEILKRYSPLRI
	ì	}	}	1	1	FHNLKDVGFLQVKVIRAEGLMAADVTGKSD
		ì	1			PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL
				1		*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS
				{		ITLIEGROLKAMDSNGLSDPYVKFRLGHQKY
						KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT
	1	1	ł	1		AWDKDAGKRDDFIGRCQVDLSALSREQTHK
	1	1	Į	1		LELQLEEGEGHLVLLVTLTASATVSISDLSVN
	İ	1	-			SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV
Ì	1	1			ł	KVIRAEGLMAADVTGKSDPFCVVELNNDRLL
						THTVYKNLNPEWNKVFTL GGPAPNSRYAEP*KNSLAMT*AHADCENYVA
125	1475	A	1556	57	509	CGGLDNICSIYNLKTREGNVRVSRELPGHTGY
				Ì		LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT
İ	1	1		1	1	TTFTGHSGDVMSLSLSPDMRTFVSGACDASS
Ì		1			1.	KLWDIRDGMCRQSFTGHVSDINAVS
1						KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL
126	1476	A	1592	3	178	EMLPTCDLADQHNIKFHYAFALNR*ER
İ		1				TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS
127	1477	A	1612	1	497	VLGAYISFGVPSSHLLTASVMSAPASLAAAKL
	1			-		FWPETEKPKITLKNAMKMESGDSGNLL*AAT
1		1	1	j		OGASSSISLVANIAVNLIAFLALLSFMNSALA
		-				WVGNMFDYPQLSFELICSYIFMPFSFMMGVE
ļ		-	- {	1	]	
L					100	WPDSFM CCMNSKAQESVFKNVLCNPPALSEMPDVKA
128	1478	A	1619	286	486	EDEVDFRASSISEEVAVGSIAATLKMKQGPM
1					1	
						PTRGALRYWIFGRFLCNIWAAVDVRCCTATI
129	1479	A	1627	1	395	MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA
			}	1	1	LLCVWALSLVIYIGPLLGWRHPAPEDETICQI
	1	1				NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV
1	1	I	}	1		•
						DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG
130	1480	A	1638	2	466	ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS
	1	1	1	ĺ		ECHNIPCEKGDKLKLFCFRLKKREIMISKLING EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHI
1		1	l			KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT
ł	)	1		1	_	KESEASTILEESHLKIPUMEETIPSSSSTIKVI
		١				KDKDIK*LLFNLYSSVEILPEVLHLK1
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVF
			1			AIRYCHGCGVAHRDLKCENALLQGFNLKLTI FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ
	1	1	- 1	1	· 1	1 FGFAKVLPKSHKELSUIFUGSIAIAAPEVLQ
	1	1	1	1	1	GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TDIPKMLWQQQKGVSFPTHLSISADCQDLLK RLLEPDMILRPSIEEVSWHPWLAST**KQWQV
132	1482	A	1656	150	48	LSNKVGGESKPKKKK  LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV VDAQ
133	1483	A	1660	3	406	RKHIKLLIQKLSDVP*ECONNQL*KLTEICEKE KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF PNFTP
134	1484	A	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP FFPAGAPPASSSSSSSSSSSSPPTVSTAPPLIPPPGF PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS SSSSSSSSSSSSSSPDRDRER*RTREREREDHS PTPSVFNSDEERYRYREYAERGYERHRASRE KEERHRERRHREKEETRHKSSRSNSRRRHESE EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE STEATPAE
135	1485	A	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL GYRYWAGIGVLQSCESALTHYRLVANHVAS DISLTGGSVVQRIRLPDEVENPGMNSGMLQE DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR GV*QNHQRAFDYFNLAA
136	1486	A	1678	525		ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS GSSSTASSLNFSAIMGSSSATASWVLSTASTPP CPSALPSSPAQES*SLAASSSAWPVAGISPSGA CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD SSSLSL
137	1487	A	1680		2999	AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE DNAELNNQNFYLSKQLDEASGANDEIVQLRS EVDHLRREITEREMQLTSKQTMEALKTTCT MLEEQVMDLEALNDELLEKERQWEAWRSVL GDEKSQFECRVRELQRMLDTEKQSRARADQ RITESRQVVELAVKEHKAEILALQQALKEQK LKAESLSDKLNDLEKKHAMLEMNARSLQQK LETERELKQRLLEEQAKLQQMDLQKNHIFR LTQGLQEALDRADLLKTERSDLEYQLENIQV LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA KKKKVPLQYNELKLALEKEKARCAELEEALQ KTRIELRSAREEAAHRKATDHPHFSTPATARQ QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST PEEFSRRLKERMHHNIPHFFNVGLNMKATKC AVCLDTVHFGRQASKCLECQVMCHPKCSTC LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT KEPSSSLHLEGWMKVPRNNKRQQQGWDRK YTVLEGSKVLIYDNEAREAGQRPVEEFELCLP DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLAPAPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCHFTNYSILIGTNK FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of nucl-	NO: of peptide	hod	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
		1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		1		to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	1	09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		(	914	ng to first		T=Threonine, V=Valine, W=Tryptophan,
	1	Į.		amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	l	l	residue of	sequence	t=1 yrosine, A=Unknown,sup count,
	ł	İ	1	peptide		/=possible nucleotide deletion, \=possible
	<u> </u>	ł	l	sequence		nucleotide insertion
	,				<b>,</b>	YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF
	1	· ·			1	NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA
		i				ISSGATYLASSYQDKLRVICCKGNLVKESGTE
	1	ĺ	1			HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	A	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
	1	1		1		PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
		1	ì			PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
	1	}	ł	1	i	CWTRGCQTTARTAAAAAAPGPAGRRPPGGA
	1	ł	1		ļ	PONGSCAASASQEAAAPPPMCPPGRRWAVAS
		ļ	1	1	l l	PPETRCPAAPGTRCRRLEAA
	<del></del>	<del> </del>	1.00	<del></del>	1276	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE
139	1489	A	1693	3	376	LISMSING SCHEDUCKTES IN AGRECULTURE
	1	1		1		FIETKALGCAWFSLCYYLVLYFESSHKVDFVF
	1	1	1	1	j	IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR
						RWAGIAKGVGTQKIIGRVHLGEQKALGL
140	1490	A	1704	3	376	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA
		i				HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL
		}		•	1	KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK
	1		· {	i		LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ
171	1471	1.	17.35	1 -	502	DKLELELVLKGSYEDTQTSFLGTASAFRFHY
	}		1			MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV
	1	}	i	1	i .	PLRPLTIVKEVTMDVPAPNVRGLNWMG
	1.400	+	1770	<u> </u>	406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS
142	1492	Α	1769	1	400	LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF
	ĺ	1		1		
	1			1		GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG
	1	1	]	1		LLQVGDRVLSINGIATEDGTMEEANQLLRDA
	l	1		L	<u> </u>	ALAHKVV
143	1493	Α	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ
	i	1	i	1	ł	KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA
	1				1	NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA
	1	}	1	1	ļ	SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL
	. 1	1	1	1	1	LENCMEMHCMDLPTDTSALVR
144	1494	A	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL
177	11,54	1.,	20	1	1	PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
1	1	ļ	1			KCNGEWVSQNDHVTQEGLDEATGLRVREVH
	1	i	1		i	IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK
i	1			1	i	SRRAYVRI
145	1,400	<del></del>	1007	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
145	1495	A	1827	20	440	CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE
	1		İ		1	
	1	1		1	1	THLALCPIVQHPEDTCIHSREVGVVCSRYTDV
1		[	1	1	1	RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
	}	_			1	PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC*
		1				SMAAET*HHVPASGADPYVRVYLLPERKWA
	1		1	1	1	CRKKTSVKRKTLEPLFDET
147	1497	A	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
1 **'	1 4777	1	1 .333	1	1	VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
						TSVTYSMG*HGAPTGSEAGANWNH**LHAH
		1.		1		YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
1		1		1		1
			<del></del>	+	_ <del> </del>	Q LLCAY DDVOCTODE A SESNA PTIVOVTA CRAG
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
}		J	}	1	1	IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
				1		GIEGRLTADQLNSATACIFAAEVAIKESERFN
		1		1	ļ	GIPALSVPVAEPIRHAEALMQQALTLKRSDET
		1		1	1	RTVQQDTQPVKSVKTELKQALLSGISFAVPLI
		- {			1	VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	A	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP
147	1477	\ \frac{1}{2}	1,000	7.1	1	PSOIRVVATATLRLAVNAGDFIAKAQEILGCP
		1	1	1	1	i
1	- }	1	1	i	- (	VQVISGEEEARLIYQGVAHTTGGADQRLVVD

					<del></del>	(
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid. E=Glutamic Acid.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	<b>\</b>	in	nucleotide	location	
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ì	i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ŀ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł		ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ			peptide	ì	/=possible nucleotide deletion, \=possible
			<u> </u>	sequence	l	nucleotide insertion
						IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
	1	1	1		ļ	YFADRNLGQENFDAAQKAAREVLRPVADEL
		1		1	] .	RYHSWKEVRGASVTVQALQEIMMAQGMDE
	1			l		RITMEIWPVD
150	1500	A	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH
				1		LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD
		1	1	1		EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS
	1	ł	1	į	į	PESSILDGMIRQLQQQQDQRMGADQDTIPRG
			1	ļ		LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV
	ŀ	ì	ł	ł		EGVRQMHQNAPRSQIATERDLQAWKRRVVV
	Ì	1	1	Į	1	PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK
	1	1	1		1	TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	A	1900	141	785	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA
	1	1	1	[		MEMOIKKOFODTCKVQTKQYKALKNHQLEV
	1	1		1	1	TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI
	İ	1	I	ſ	{	NEMMASQALRLDEAQEAECQALRLQLQQEM
		ì		1		ELLNAYOSKIKMOTEAOHERELQKLEQRVSL
	1	1	ì		İ	RRAHLEOKIEEELAALOKERSERIKNLLERQE
İ		1				REIETFDMESLRMGFGNLVTLDFPKEDYR -
152	1502	A	1915	12	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL
132	1302	^	1713	-	13,,	RQKIFKERALPDIENYMFENHDQLRQAATEC
	1	1	1	1	Į.	MCNMVLHKEVQERFLADGNDRLKLVVLLCG
		,	ŀ			EDDDKVQNAAAGALAMLTAAHKKLCLKMT
}	1	1	1	1	į	OVTT
L	1503	+	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL
153	1503	A	1921	1	237	ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG
		1	1	<b>\</b>		FNFLLYMIFLYT
	<del> </del>	<del></del>	1000	<del> </del> 2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR
154	1504	A	1928	1 2	334	YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV
1			ı			RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS
	1	1	1			
	<del> </del>	<del> </del>	1.000	<del></del>	1.00	QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKYEKELTMFQNDFEKACQA
ļ	ł	1	1	1	ļ	KSEALVLREKSTLERIHKHQEIETKEIYAQRQ
l		1	1		1	LLLKDMDLLRGREAELKQRVEAFESYQLELK
i	J					DDYIIRTYRLIEDDRINIQISGHWQESP
156	1506	Α	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENEL
1	ì		ľ	1	1	DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ
						RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	A	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA
1	1 .	1	1		1	AYEYYDAGNHWCKDCNTICGTMFDFFTHMH
1			_L	1	<u> </u>	NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	A	1939	1	423	TTHRLNYTAEPPCTSMPIYWMPDVPHRCTTA
1	1	1		}		NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED
1			1			RLHCQTQACPPLSWPQRLDILLGTARAIQFLH
		-	}		1	QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA
						RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR
1.00	1.557	1	1 - 7 - 7	1		NGHLATVKLLVEEK ADVLARGPLNQTALHL
			1		1	AAAHGHSEVVEELVSADVIDLFDEQGLSALH
1	-		1	1		LAAQGRHAQTVETLLRHGAHINLQSLKFQGG
ì	1	1		1	1	HGPAATLLR
160	1510	A	1982	12	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA
100	. 1310	1^	1702	-	1711	IYSEYCNNHPGACLELANLMKQGKYRHFFEA
1	1	- 1	1	1		CRLLOOMIDIAIDGFLLTPVQKICKYPLQLAEL
1	)	1			1	LKYTTQEHGDYSNIKAAYEAMKNVACLINER
1			.		1	KRKLESIDKIA
				1	1	
-	<del></del>	→	1004	<del>-</del>	770	DETACKSI SDEGI ECAESVAVIONI VSSDNIVE
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR SSHNASKTLNANNMETLIECQSEGDIKEHPLL

		·	CODO	75 11-2-3	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		i	!	peptide	{	/=possible nucleotide deletion, \=possible
		Ĺ		sequence		nucleotide insertion  ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
			Ì	l .		PASDFSGALETDLKASLFDQPLSIICGDSDTLP
			ì		1	RPIQDILTILCLKGPSTEGIFRRAANEKARKEL
	}	}	ł	ł		KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
	ł	ŀ		ļ		RKLLSSDLFEEWMGALEMQDEEDRIEALK
1/0	1512	A	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
162	1512	A	1900	1 804	301	KFQGRWGTVCDDNFNIDHASVICRQLECGSA
	]	ļ	}		1	VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
	1	}	1			CKHOGWGKHNCDHAEDAGVICSSKD
163	1513	A	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD
103	13.5	1				LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
		1				FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG
	1		1	1		SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
			1	}		ENNWYFVVADSSKAGFTTIYKWERETGFYSH
						QSFTR   EDPEELGHFYDYPMALFSTFELFLTIIDGPANY
165	1515	Α	2013	2	403	NVDLPFMYSITYAAFAIIATLLMLNLIAMMG
	İ	i		<del> </del> -		DTHWRVAHERDELWRAQIVATTVMLERKLP
			•	1		RCLWPRSGICGREYGLGDRWILRVEDRQDLN
			1	1	(	RQRIQRYA
2//	1516	+A-	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
166	1516	A	2019	*	121	QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
				1		NKNMVTERLONVMVLEQCFSDSSSLYRFLTY
	1	1	1	ì	1	SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI
		1	1 "	<b> </b>	1	WDMRNLATIFLAVVMALLSLHCLAAFKRLE
	1		1	1		HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
		ł	1	}		VLYMPSMGYCILFVHGLSKLCTWLNRCGATT
	1	ļ		l		LIVSTVLLLLFSWKTVKQNEIWLSRESLFRS
1		j	}	1		GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
l		<u> </u>				HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW
ĺ						PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY
	1		i		}	GRLFAVVHFASRQWKVTSEDLILIGNELDLA
l		ĺ	1	1	1	CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
1	İ	1	1			RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
Ì	i	ł	1			Levanti vizzien zu internet in
						TTPOTVLRINSIEIAPCLL
169	1512		2046	12	366	TTPQTVLRINSIEIAPCLL HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
168	1518	A	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLOGAARVFMPLQAQVKAKASKPLQMQIKA
168	1518	A	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
168	1518	A	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
		A	2046	2	366 945	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR ONLEDREVLNGVQTELLTSPRTKDTLSDMTR
168	1518					HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVV KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKNNVDLVDKTFA OAODVFROAMKSPSVLLHVLPPQNREQYEKS
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
169	1519	A	2049	1	945	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVY KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH
						HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVY KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF
169	1519	A	2049	1	945	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE
169	1519	A	2049	1	945	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
169	1519	A	2049	363	945	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA
169	1519	A	2049	1	945	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV

NO. of   No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
			,				D=Aspartic Acid, E=Glutamic Acid,
USSN   Ostation   Os				in		location	F=Phenylalanine, G=Glycine, H=Histidine,
19496	eotide	,	1	USSN	location	corresponding	
1914   ng to first amino eidd residue of peptide peptide peptide peptide peptide peptide sequence   Particular of peptide peptide sequence   Particular of peptide peptide sequence   Particular of peptide sequence   Particular of Peptide peptide sequence   Particular of Peptide sequence   Particular of Peptide Sequence   Partic	seq-			09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
residue of peptide   pep			İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
peptide   sequence		Ì			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	}	]	}		peptide		/=possible nucleotide deletion, \=possible
RESCTITATYCOSYODYNNOKCVSGGTSQKE	ł				1	ļ	nucleotide insertion
SCPLLGYYADMWKDILLGKRUPPMTKAFFDT				<u> </u>			RESCTITAYPCDSYQDYRNGKCVSCGTSQKE
172	1		ĺ				
	ł		ł			1	AEESPFCMYHYFVDIITWNKNVR
	172	1522	A	2056	3	361	LIQHKSAVEYAQSHLSLVSMCKESHKCSEPK
HLVRGKEQRRREEMMRILKCLKES		İ		į		<u> </u>	MEWKVKIRSDGTRYITKRPVRDRILKERALKI
173	1	ļ	}	]		]	KEERSGLTTDDDTMSEMKMGRYWSKEERKQ
174			ļ		1		HLVRGKEQRRRREFMMRIRLKCLKES
AVFLSYILTYTGYIAPWSGRFYSLUDTGYA	173	1523	Α	2060	1	387	GTRILSMQIPFVGFQPIRTSEHMAAAGVFALL
174	1		Ì		1	ļ	QAYAFLQYLRDRLTKQEFQTLFFLGVSLAAG
174	}	ŀ		1	ļ	ì	
174		į	1		•	i	KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA
RRILEREERARILYURKEEMBRILEIQRIRKEKW   HRILEAKDLERRNEELEIYILERCFPEAEKLK   QETKLLSQWKHYIQCDGSPDFSVAQEMNT   1525   A 2083   139   486   AALTWSQRQEFWRMEMQPIVTDMVTHWV   AESSTYGWLCALFRVTHVOGATGHGVVCG   RRVLCGLPLPSPAPMPIMSLPEGESSKEREVQ   RLQFPYLEPGHELPATILAFLAAV     176					1		,
HRLEAKDLERRNELEELYLLERCFPEAEKLK	174	1524	A	2071	74	443	
1525   A 2083   139	1	ľ		1	1		
175		į.	ļ	1			
AESSTVGWLCALFRVTHVGVGATCHGVVCG   RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ   RQFPYLPSPAPMPIMSLPEGESRKEREVQ   RQFPYLPSPAPMPIMSLPEGESRKEREVQ   RQFPYLPSPAPMPIMSLPEGESRKEREVQ   RQFPYLPSPAPMPIMSLPEGESRKEREVQ   RQFPYLPSPAPMPIMSLPGESSPAFKPSMRS   FOKFLMLLGDTITLKGWTGYRGGLDTKNDTT   GHSVYTVYQGHEIMFHVSTMLPYSKENKQQ   VERKRHIGNDIVTIVFQEGEESSPAFKPSMRS   HFTHFALVRYNQQNDNYRLKIFSEESVPLFG   PPLPTPFVFTDHQEFRDFLLVKLINGEKATLET   PCI   GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC   CDGAWLAWACWVFGNDFSPASAACSALLG   CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR   RAVSVPLTLAETVASLWPALQELARCGNLAC   RSDLQ   RAVSVPLTLAETVASLWPALQELARCGNLAC   RSDLQ   RAVSVPLTLAETVASLWYAESKDQVFQ   PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF   PCCAPPILLVKGSKPSQQGRYNNTVERFSSSL   LILQVREADAAVYYCAVEVPNTDKLIFGTGT   RQVPFNIQNPD   LILQVREADAAVYYCAVEVPNTDKLIFGTGT   RQVPFNIQNPD   RTRSSTRPSSLFVHASAKGGEKEEGDDGHYL   MRTESHTGLKKGGNANLVFMLKRNTEPKKG   SYHFDLERLRAAHLEREQEHLAPGGISMLP   PPPLPLACLG   TSKRABETTDVTRSFGWDSSEAWQQHDVGE   LCRVMFDALEQKWKQTEQADLINELYQGKL   KDYVRSLECGYEGWRIDTYLDJPLVIRPYGSS   QAFASVVCTFHLTJEREQGELAPGGISMPL   PPPLPLACLG   STREABETTDVTRSFGWDSSEAWQQHDVGE   LCRVMFDALEQKWKQTEQADLINELYQGKL   KDYVRSLECGYEGWRIDTYLDJPLVIRPYGSS   QAFASVVCTFHLTSCYLDIPLVIRPYGSS   QAFASVVCTFHTLTSVTV   SSAVVCTFHLTSVTVV   SWVTSYLVMVSNDSHTWVTGKNGSGDMIFE   GNSEKBIPVLNELPVPMVARYTRINPQSWFDN   GSICI   GSICI   GWVLGKIMCKITSALYTLIVVSGMQFLACISI   DRVVAVTKVPSQSGVGKPCWICFCVWMAAI   LISIPQLVFYTVNDNARCIPIFPRYLGTSMKAL   LQMLEICIGFVVPFLIMGVCYFITATILMKAL   LQMLEICIGFVVPFLIMGVCYFITATILMKAL   LQMLEICIGFVVPFLIMGVCYFITATILMKAL   LQMLEICIGFVVPFLIMGVCYFITATILMKAL   LQMLEICIGFVVPFLIMGVCYFITATILMKAL   LQMLEICIGFVVPFLIMGVCYFITATILMIAVC	L		<u> </u>		1		
RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ   RLQFPYLBPGHELPATTLLAFLAAV     176	175	1525	A	2083	139	486	
RI.QFPYLEPGHELPATTILAFLAAV	ĺ	ſ	1		1		
176	1			1 .			
FQKPLNLLGDTITLKGWTGYRGGLDTKNDTT   GIHSVTYQGHEIMFHVSTMLPYSKENKQQ   VERKRHGNDIVTIVPQGEESSPAFKPSMIRS   HFHHIPALVRYNQQNDNYRLKIFSEESVPLFG   PPLPTPVFTDHQEFGESSPAFKPSMIRS   HFHHIPALVRYNQQNDNYRLKIFSEESVPLFG   PPLPTPVFTDHQEFGDFLLVKLINGEKATLET   PCI   GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC   CDG AWLAWACWVFGNPF9PASAACSALLG   CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR   RAVSVPLTLAETVASLWPALQELARCGNLAC   RSDLQ   RSDLQ   RSDLQ   RAVSVPLTLAETVASLWPALQELARCGNLAC   RSDLQ   R				·	<u> </u>		
GIBSVYTVYQGHEIMFHVSTMLPYSKENKQQ	176	1526	A	2092	3	587	
VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS	1		1				
HFTHIFAL VRYNQQNDNYRLKIFSEESVPLFG   PPLPTPYFTDHQEFROFILLVKLINGEKATLET   PCI	1		1		i		
PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET   PCI	1	,			Į.		
PCI		1		Į.		ļ	
177		1	1				
CDGAWLAWACWVFGNDFPSPASAACSALLG		1	ļ.,	L		100	
CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR	177	1527	A	2103	44	427	
RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ  178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD  179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHLLFEREQEHLAPGGISMPL PPPLPPACLG  180 1530 A 2116 3 366 TISKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLEGCYEGWRIDTYLDIPLVTRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV  181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFFGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWILGFCVWMAAI LLSIPQLVFYTVNDNARCPIPFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1						
RSDLQ		)	]			ļ	
178	1		1				-
PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRILVKGSKPSQGGRYNMTYERFSSSL LILQVFPNIQNPD  179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV  181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	170	1500	<del>                                     </del>	0104	1	400	
PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD  179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG  180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHINSTVV  181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDYJILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	178	1528	A	2104	2	409	
LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD  179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG  180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV  181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWYTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1		ļ		}		
RLQVFPNIQNPD   1529   A   2111   1   312   PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL   MRTESHTGLKKGGNANLVFMLKRNTEPKKG   SYHFDLERLRAAHILFEREQEHLAPGGISMPL   PPPLPLPACLG     180			1				
1529   A   2111   1   312   PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL   MRTESHTGLKKGGNANLVFMLKRNTEPKKG   SYHFDLERLRAHILFEREQEHLAPGGISMPL   PPPLPLPACLG     180			1		1		
MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLACLG	170	1520	<u> </u>	2111	<del>                                     </del>	312	
SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG  180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV  181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1''	1329	1	2111	1.	312	1
PPPLPLACLG   PPPLPLACLG     PPPLPLACLG     PPPLPLACLG     PPPLPLACLG     PPPLPLACLG     PPPLPLACLG     PPPLPLACLG     PPPLPLACLG   PP	}		}				
180   1530   A   2116   3   366   TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV     181		1			1		
LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV  181 1531 A 2117 2 386 YGLGAHFGRLFJQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	180	1530	A	2116	3	366	
KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV     181	1	1550	1"	1 -1.15	1		
QAFASVVCTFHLTACVSLHRIHNSTVV	1	1	}		ļ		
181   1531   A   2117   2   386   YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI   182   1532   A   2123   1   493   RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS   1533   A   2140   3   561   RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			ŀ	1			
NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	181	1531	A	2117	2	386	
DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMGFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1		1		1	300.	
GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI  182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1	1	1	1		
IS2   1532   A   2123   1   493   RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEIGFVVPFLIMGVCYFITARTLMKMP NIKIS   1533   A   2140   3   561   RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1		ľ	l		
182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			1		1		
GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	182	1532	A	2123	1	493	
DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1		1	1	1		
LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1		1		1		1
IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS  183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF				1	1		1
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	ł	1	1		1		
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1		1				
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ
VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI			1		1		
	L		1		<u> </u>		VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

			000	Designation	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQID	SEQ ID	Met	SEQ ID NO:	Predicted beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid,
NO: of	NO: of	hod	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
otide	seq- uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
eq-	uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
ience			'	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
•				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	·	/-possible nucleotide deletion, \-possible
	<b>'</b>			sequence		nucleotide insertion
		<del> </del>	<del> </del>			SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS
	1	[	1		<b>\</b>	EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR
				ĺ		RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND
104	1331	1		1		NPPVFTRASYRVTVPEDTPVGAELLHVEASD
		ļ	1			ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR
	ľ	İ	ļ	ļ		LAHALDCETQARHQLVVQAADPAGAHFALA
		1	1	1	1	PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG
		1		ł _		TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW
.05		1				PKNNFNGSLVQASYQHEELRREVIMLACSFG
	1	1	1		1	NKHCHQQASTLISDWISSNRNRIPLNVRDIVY
	1	1		1	,	CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL
	1	1			1	LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI
	1	1	ł	1		DVIIHVARNPHGRDLAWKFFRDKWKILNTRI
			}			RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH
		1			<u> </u>	E CONTRACTOR WETTER
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYIVHPLWETWA
100	1.550		1	1	•	HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP
	1	1			'	LDEQNRDWQGLLENLHVELTLDEEDSEGPEK
	}	1	1		1	EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC
	İ	1	ı			AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR
10,			1	Ì	1	PGAVAYTCNPSTLGGWGGWITRSGVRDQPG
1		1	i			QHGGTPS CONTAINE AND AND AND AND AND AND AND AND AND AND
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA
100	1000	"-	1	İ	ļ	CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP
	ŀ	1		1	Ì	AVVVPYMMVLQENGYGVEEGIPTLLMAASS
ł			İ	1	ŀ	MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS
}		ł	1	1		LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG
ì		-		1		HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF
1		ł	l	1		QALAPDLINGURKLI I LHIMRANAIQI VI VIGI QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE
		1		1		LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV
		- 1	ł	1		LHLEHNDLVKVNFANFIKLISLINSECERGGIEV
		1	<u> </u>	l		AIVVSSLDW MRLNQNTLLLESFGXXRPYTSEHAPTYHQW
190	1540	A	2179	64	399	MRLNQNILLLESPOXXRPIISEMAPIIMQW
1.70		- [	-	1	i e	MKADELLRWTTSEPLTLEHEYAMQRTWLED
		- 1	1			AYECTFIVLDAEKRHAQPGATEESCMVGDVN
1		1				LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR
***	120.12	1 -		i i		LSYVLFIQERDVHKGMFATNVTENVLNSSRV
	1	- 1	1	l l		QEAIAEVAAELNPDGSAQQQSKAVNKVKKK
1			- 1	- [		AKRILQEMVATVSPAMIRLTGWVLLKLFNSF
		1	-	1	1	FWNIQIHKGQLEMVKAATETNLPLLFLPVHR
		1	1			SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR
172	1.5.2	1				DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS
193	1,343	1				YTHSKGIMHRDVKPLNILCNSPRNKVILADW
1	1	1	1		1	GLAEFYHPMRKYSVHVATRYYKSPEILLDYE
1			1			YYDYSLDIWAVGVILLELLTLKLHVFEGGDN
ł	l	- {	1		1	EQ
194	1544		2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMG
194	1344	1 ^		1 - 7 -		NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE
}	1	- 1		1		LPVPMGARYIRINPQSWFDNGSICMRMEILG
1	1	1				PLPDPNNY
195	1545	-	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETC GAVSSLQIVTELQTNYIGKGCDRETYSEKSLC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrossile, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW MKHGPSPGVRAEKETILCYSDKTEMNRHHY ALYVHNCRLVFLLRKDFDQADTFRPAEFHW KLDQQALAKVDGQPGKSITRQLQEMPVTIQG ISLKPS
196	1546	A	2256	1	396	FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF KALEESGALLESGTYDGNFYGTPKPPAEPSPF QPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL HCFTESGRGKSWRLCAAILPL
198	1548	A	2275	3	404	TCITVVVIPRMLVDFLSESKTISLPECATQMFF FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT LMTRKICLQMMMASWMVGFLFSLCIIVTVFN LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM AIFVLSA
199	1549	A	2315	Î.	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI KRLAFCHSNVLSHSYCVHQDVMKLAYADTL PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH DQLQFEYKEEM
201	1551	A	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEEGAGHIKDLYL LIMKDESLYQGLREDTLRHQLVETVELKIPE ENQPPSKQVKPLFRHFRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T
205	1555	A	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC EDCSCR

CE() II)	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=A spartic Acid. E=Glutamic Acid.
nucl-	peptide	1.00	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
neuce	ucitoo		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		}	l	peptide		/=possible nucleotide deletion, \=possible
	ł	ĺ	1	sequence	1	nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH
200	1550	^	2.00			THIPCPVSSTTNDTPDQIWVSVGSLRMGTGG
	i			1		MGANASTSPRCWDLSSGNKKWIIQVPILASIV
	1	į .		į.	1	ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	A	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT
207	1337	1 ~	2407	200	1	HTIMVLDKENTLS
208	1558	la -	2413	64	492	VOGTGXXFIAFTEAMTHFPASPVWAGMFFL
200	1336	^	2413		1 .,,	MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT
		İ	1	ļ	i	GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA
		1			1	TLPLTLIVILENIAVAWIYGTKKFMQELTEML
		l .		1	1	GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP
209	1339	^	2417	1	• · ·	NASNLDKVLTDIKADKDQANDGLSSALLILY
ļ	)	Į	1			LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ
1		1	ļ	l .		DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM
	1	ł	-			TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE
	1	ļ	ì	1	Į.	RPPDHOHSAOVKRPSVSKEGRKTSIKSHMSG
	1 .	1	<b> </b>	1	1	SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE
		ļ	1	Ì	1	AGPOGLHDLGRSSSSLLASPGHISVKEPTPSIA
}	1	1		1	1	SDISLPIATQELRQRLRQLENGTTLGQSPLGQI
		-		1	İ	OLTIP
210	1560	TA.	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA
210	1300	1 ^	2422	"	1	AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW
}	1	1		į		GAVRAAESLTDIAEPASPQVHETPIDASQTQK
		ļ	ł			VEPASKSRFTPELQAKVSHSRERALSTMDATP
	1	1	ł	Ì		HHAOPORGEG
211	1561	A	2431	1	764	RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP
211	1301	1 **		1		LMFWLHGIQLVALNYQTDDLPLHLNAAMFE
	1	1	1	i.		ANGGCGYVLKPPVLWDKNCPMYQKFSPLER
	1	1			ł	DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE
		1				VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF
1	Ì	}		}		LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL
1		-	1			KALKRGYRHLQLRNLHNEVLEISSLFINSRRM
1	1	1	1			EENSSGNTMSASSMFNTEERKCLQTHRVTVH
		1		1		GVPG
212	1562	A	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI
		1	1	1		YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC
1		1			1	TDLDHNLDKGYLTVLGEQATPTNRLGALPKG
1						RANKTRDLELTYLAERIVRLTWIPGDANNRPI
		j		j		TDYDCQIEEHQ
213	1563	A	2445	1	1294	MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT
	1	1		}		WLMPELHPKEQILELLVLEQFLSILPEELQIWV
}		1	İ	İ		QQHNPESGEESVTLLEDLEREFDDPGQQVPAS
		į	ł			PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ
1	]	Į	-	1	1	LKSWKPCLSPKSDCENSETATKEGISEEKSQG
	1	- [	1			LPQEPSFRGISEHESNLVWKQGSATGEKLRSP
			-	1		SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT
1				J	1	TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT
		1	1	1	1	QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR
)		1	1	ł	1	IHSGEKAYECSECGKAFNQSSALIRHRKIHTG
		ł	- 1		1	EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY
		- 1	1			ECNECGKTFSQSSKLIRHQRIHTGERPYECNE
	}	1	1			CGKAFRQSSELITHQRIHSGEKPYECSECGKA
		1	1			FSLSSNLIRHQRIHSG
1			2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL
214	1564	l A				
214	1564	A	2.01	1		STRLVSVQEDAGKSPARNRSASITNLSLDRSG
214	1564		2.0.		·	SPMVPSYETSVSPQANRTYVRTETTEDERKIL
214	1564		2101			STRLVSVQEDAGKSPARNRSASI I NLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalamine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciico	ŀ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence	ľ	[	///	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ì	Ì	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	l	1	peptide	Soquence	/=possible nucleotide deletion, \=possible
	ł	1	[		1	nucleotide insertion
		<u> </u>	<u> </u>	sequence	ļ	AMVDGRWLDCVSHHDQLFRDEYALYRPLQV
		1	1		Ì	
				L		LFSVYCQLECSKLIL
215	1565	A	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ
				1	į.	HGPGRHGRRVCSSQDSMADVFVHLRTAWPT
	l	i	1	ł	ì	CSLISGQHGPGESVSYEDDDIPAPASLLHVNA
	Į.	1	ſ	1	ì	AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR
	į.	}	1	i	ļ	QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
	1	1	1	l	}	TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
	1	1	1	1		STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1		1	1 .		TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
	1	Ì	1	)	}	ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
		1	į.		İ	
		1		1		TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1		,	J	}	AIVEGVKISSCIPDLICAVSIESI VEGVKISSK
		[	1	1	-	TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
	1	i	1	1		ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC
	1	1	1		1	TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
	ì	1	1	1		ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1	1	1		1	TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
	1		1	l		STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
}	}	}	4		1 -	TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
	ì	1	1		}	STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
[	1		1		1	PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS
}	1	1	j		1	
1	l l	1				TVPGVHISSCTTDLTCAVSIHATVPGVHISSCT
ł		1	i	1		PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHS
	1	1	1			TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT
ľ	ł	1	1			PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA
Į.	1		1	1	1	TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT
1	1	1	ľ	1		PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
ł	ı	1	-{	1	İ	TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
Ì	1		1	1		PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
	1	1		1		STVPGVHISSRTPDLTCAVSIHATVPSVHISSC
1	1	1	1	j	}	TO A VOID TO THE TOTAL OF THE STREET
		1				TPDLTCAVSIHSTVPGLLTSVSQTSTG
216	1566	A	2477	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK
1	1	1	1	1	1	AVEVATVVIQPTVLRAAVPKNVSVAEGKELD
i	1	1		Ì		LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
1	1	1		1	1	RVLARLDRDFLVHSSPHVALSHVDARSYHLL
1	1	1		}		VRDVSKENSGYYY
217	1507	A	2480	12	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
217	1567	A	2400	-	100	MQVLVCQHECVRELATRPGRLSPIENFLPLHY
1	1				1	DYLQFAYYRVGEYVKALECAKAYLLCHPDD
1	1			i		EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
			1	1		
		_L_				VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
} _		Į				SANLLQLVRSSGDIQEGDLVEVVLSASATFED
		1				LQIRPHALTVHSYRAP
210	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLQ
219	1303	A	2409	1	120	CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
1			1	1		DCKYECMWVTVGLYLQEGHKVPQFHGKWP
1 .	1	1		1	1	
1	1	ſ		[	1	FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
		1				FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
1		1	1	1	(	APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
		ı	-			HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
	1	ſ		1	1	MLKCRVDNVNSQLQVLGDHLGNTNADIQMV
1	}				1	KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
		1	ı	1	1	KEDLEKADALTFQTLNFLKSSLENTSIELHVL
1	- 1	- 1	- 1			T K EULEKADAL LEULINGIAGALEMIGIGIGUT
		ł	Ì	1		
						SRGLENANSEIQMLNASLETANTQAQLANSS

CTO TO 1	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			·	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ	Ì	peptide	ļ	/=possible nucleotide deletion, \=possible
		ļ	}	sequence		nucleotide insertion
		1				FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
	1	1				VTAQTQKANGRLDQTDTQIQVFKSEMENVN
	}	)		}	Ì	TLNAQIQVLNGHMKNASREIQTLKQGMKNA
	1	1	1	•		SALTSQTQMLDSNLQKASAEIQRLRGDLENT
	<u> </u>	\		<u> </u>		KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	1571	A	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFQHIIRREV
		1	l	l .		TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD
	ł	1	ł	ţ		TCGEEASVLEILVYNSKIENRHEMLAVEPINE
		ì	1			LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT
	1		Ì	1		GVLFFFTN
				<u> </u>	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
222	1572	A	2508	3	393	ARHVRDKEEEVDLVMQKVESLRQELRRTER
·	ľ	1	Ì	Ì		AKKELEVHTEALAAEASKDRKLREQSEHYSK
		1		i	1	QLENELEGLKQKQISYSPGVCSIEHQQEITKL
		1		ł		KTDLEKKS
223	1573	A	2544	+2	412	NDPAIISNFSAAVVHTIVNETLESMTSLEVTK
223	13/3	Α.	2544	~	112	MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
l		1		}	:	EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
		1				KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ
			Ì	l		LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
224	13/4	1"		1	-	CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
1	İ	Ì			1	QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL
ļ	1	1		İ	j	TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
		ĺ	1	1		VK .
225	1575	A	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
	10.0	1		1		GLGEEEEKEAGKKKKKQEEKEKEKGAVYSR
1		1		1	1	VARICKNDMGGSQRVLEKHWTSFLKARLNC
	1			1	1	SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
		1				QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP
			-			DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
	}		1	}		TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
	-					KTRVRYRLTAISVDHSAGPYH
226	1576	A	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP
j	ļ		1		1 .	TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
	1		1			AGAACDRGMSLEACEAVTRKANRRTYTMG
		1		1	Ì	VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
L					1105	RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
1		1		1	1	EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
1			1	ļ		ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE
	1	1	1	1		KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
			1	1		PSEYESGSACPAOTVHYRPINLSSSENKIVNVS
1		- 1		1		LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ
1 .		- [	1	1		LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS
1		1	1			HRAQIMARYPEGYRTLPRNSKTRPESICSVTP
		1		1		STHDKTLGPGAEEKRSMRDDTMWQLYEW
				1		QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT
	}	1	-			MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
		1		1		ORGDVTIDRRHRAHHPKVK
					220	LPFLGLGSVLPQGMVMASPEMNPTICSVFEA
228	1578	A	2583	3	330	HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
		1	1	1		KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD
l		i		1	•	PTMGIKPHLWWVAA
	2		0.505	<del></del>	140	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR
229	1579	A	2589	1	448	ECVAPNICKCKPGYIGSNCQTALCDPDCKNH
ł		1		1	į.	GKCIKPNICQCLPGHGGATCDEEHCNPPCQH
1	- 1	1	1			ONCING MICKORI GITGOVI CONDITION LOCAL

		-		<b>*</b>	Donalist Jane	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted beginning	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
cotide seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
neuce	uence	}	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
ucace			7.4	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		}	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide	'	/=possible nucleotide deletion, \=possible
	ĺ	]	1	sequence		nucleotide insertion
			<u> </u>			GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
,			1			ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	A	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI
						LYLLYLLRNAFFLR
231	1581	A	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
		<u> </u>				WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
		1				GRARRTPTCEPATPLCCRRDHYVNFQELGW
	1			]		RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
	1	ļ				LLYL
				<u> </u>	102	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
233	1583	A	2601	184	403	YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
	1			1		NSIPYWERIT
	1.204	<del> </del>	12614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
234	1584	A	2614	1/8	333	DGSKSSDDQKIISYLWEKTQ
235	1585	A	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
, 235	1363	^	2010	<b> </b>	450	AKFLNVEAAMVFGMGFATNSMNIPALVGKG
	1 .				<b>\</b>	CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
		i		1		OSLEKLLRDAVIYGOPRTRRAWKKILILVEGV
		1				YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
		1	1	[		GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
					İ	FGASGGYIAGRKARILSPPACLVPNTGSHSLH
	[	1		į		RLTRDLQMNEAMVALVTDRLQGWNSGEGN
	1		}	1		WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
		1				AEQIIRSLKLIMGLDGTTQ
236	1586	A	2621	1	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
	1	İ	1	1		ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
			1	1		WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
				ļ		KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
			1		<del></del>	A DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
237	1587	Α	2628	398	1	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
		i	1	1		
		1	1	1		A CEDDOWNMSGODOR AVEDGERLEPPRICEN
1		1	1	ł		GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
		}			}	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
220	1500		2631	<u></u>	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKYYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS
238	1588	A	2631	1	1104	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN
238	1588	A	2631	1	678	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKYYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIPSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMPSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGF KEALQTHKHQNGIPSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKROLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGF KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE
239	1589	A	2636	1	678	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSG LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE KVNRDCI ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
239	1589	A	2636	1	678	LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV  WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHITFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE

000 700	COECID	1404	CEO.	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=A spartic Acid. E=Glutamic Acid.
NO: of		noa	in No.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
otide	seq-		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	Ì			acid residue	Q=Glutamine, R=Arginine, S=Serine,
ience	İ	1	914	ng to first amino acid	of peptide	T-Threonine, V=Valine, W=Tryptophan,
	İ	ļ.		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
					sequence	/=possible nucleotide deletion, \=possible
	l	1	1	peptide		nucleotide insertion
		ļ	<u> </u>	sequence		EMANKELKQLRASYTESCIQEHYLPQVIDGTL
		]			1	EMANKELKQLKAS I IESCIQENI LIQVIDGIE
		<del>                                     </del>	12640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD
241	1591	Α	2640	392	3	TCDLCGYNQKLYPCWETQVGQEMYKLMIFD
	1	Į	ì	1	l .	FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ
	1	1 .	1	1		QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM
	j			1	ł.	Y
0.10	1592	A	2642	405	+1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR
242	1592	\ \ \	2042	403	1.	MVI.MAGAKPGMLLLCFMCCTTLLSMWLSNT
	1	1		1		STTAMVMPIVEAVLQELVSAEDEQLVAGNSN
	ı	1	)	1	ł	TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK
	i	1	1	1		SPLMISQACI
		4	-	<del>                                     </del>	<del> </del>	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF
243	1593	Α	2646	412	2	LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ
	1	1	1		1	VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL
			1			VSFINFFTSVLATLVVFAVLGFKANVINEKCIT
		1	j	1		
		1				QNSETV MTTTLIGLLKTARLLRLVRVARKLDRYSEYG
244	1594	A	2650	1	1271	MTTTLIGLLK I ARLLINLYR VARALDR I SET O
•	1		1			AAVLMLLMCIFALIAHWLACIWYAIGNVERP
	1					YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK
			1			DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF
	1	1	l l	1	-	SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY
	į.		1	ĺ		HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA
	1	İ	ì	1		WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS
	1	Ì	1			NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS
	4	- [	- {		1	TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS
ŀ	Į.		i i		1	PPSSDKTIJAPKVKDRTHNVTEKVTQVLSLGA
	-	Ì		1	İ	DVLPEYKLOAPRINKFTILHYSPFKAVWDWLI
<b>!</b>	l	l	1	1	1	LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY
}	ŀ	1	1	1		SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE
!	1	1				VVSDPASV
	<del> </del>		2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW
245	1595	A	2656	363	1 -	WESLLLTAYFCYVVFMKFNVQVEKWVKQ
	1	}	- 1	1		MINRNKVVKVTAPEAQAKPSAARDKDEPTLP
	- }	- 1	ì	ł	Ì	AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD
İ			1	<b>!</b>	1	PHV
				1	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGP
246	1596	A	2660	200	306	LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG
	1	+	1		l l	LLDWKIKHLIK I IQLSQATALAKLESDIT EIG
İ	1	- 1	١.	1	1	IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS
1		{				NYFLS
247	1597	A	2678	3	267	DAWYKNDIIFNQTERKQKISENLKHLASVRV
1					1	VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF
1		1				VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	A	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF
1		1	1	1	1	GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQI
1	}	ł		1		ENDAVAEESRQVLTICAQFLKWKLPREVYSK
1		l				DPWHIKPTEAGTICRFFEKKCKGKINILEQTL
{	1		-	- 1		MYSKNPKL
249	1599	A	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV
447	1333	1	-3/2	1.	1	SFPIGIYKVLRNVSGOIHLITLANNELKSLTSK
1		- 1	1	į		FMTTFSOLRELHLEGNFLHRLPSEVSALQHLI
}	1	Į	1			AIDLSRNQFQDFPEQLTALPALETINLEENEIV
1		ı			l l	DVPVEKLAAMPALRSINL
				450	121	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRR
250	1600	A	2693	459	21	PPLGALSQALTFLSRAAKNHSQDPGKGTKPF
1	1	{	I	ŀ	ļ	AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPT
1	1	l l	ĺ	1		AAPAAPPRASEPAPEMULADAUTQFAFFI
i		- 1	- 1	}	1	NSPWHPATLPGALGPQLSQAAPSPIPPPCLM
1	i	,				
		- [			404	ISSCPDLKLTKSSTP FVFDLKLRVPGFAALLIHGASSVPGPETVRLI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ		Ì	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	l	]	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
		[		ļ		QKRKKKAPDHSSGRKEELVTTHTVDKLETKK
		}	1	]	)	PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI
		]			l	GAGDCL
272	1600	<u> </u>	2697	421	1	POKSHSGAYQCFATRKAQTAQDFAIIALEDG
252	1602	Α	2097	421	1 1	TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
	į	j		ł		VTWALDDEPIVRDGSHRTNQYTMSDGTTISH
		1				MNVTGPQIRDGGVYRCTARNLVGSAEYQARI
		}	ļ	}		NVRGPPSIRAMRNIT
253	1603	A	2698	65	401	ACCOWRRTLIPAKSTTVSCTISTPHHPFRGSYS
233	1005	) <b>^</b>	2000	05	''-	FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL
	1		1	1	Ì	PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
				1	j	AAFTNNIASSTIIL
254	1604	A	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ
-5.						RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
		1			1	AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
		1		i .	İ	DKETLENMMQRHEEEAHEKGKILSEQKAMIN
	ŀ	ł		:	i	AMDSKIRSLEQRIVELSEANKLAANSSLFTQR
		1			1	NMKAQEEMISELRQQKFYLETQAGKLEAQN
	}		İ		1	RKLEEQLEKISHQDHSDKNRLLELETRLREVS
	ĺ		1			LEHEEQKLELKRQLTELQLSLQERESQLTALQ
		1		1	1	AARAALESQLRQAKTELEETTAEAEEEIQALT
	1	<u> </u>	-		105	VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	A	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE
		1	Ì			LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA
	i	1		1		FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
		1				PYEARMFFEGLR
257	1607	A	2702	12	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
231	1007	1 ^	2,02	1	1 373	FNPSFSFLDPRYSVGGDENIGTVTTLANILREF
	į					NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
			1			DLPVQARRLVDLMKNDTRIHFQEDWKJITLFI
	İ	1	ļ		1	GGNDL
258	1608	A	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI
		1		1		MTRKELLTVYSSEDGSEEFETIVLKALVKACG
	'		į.		ţ	SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
Ì						QWRSFHLEASLMDALLNDRPEFVRLLISHGLS
	1	Ì			]	LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH
1		1	1	1		SAGTKAPALKGGAAELRPPDVGHVLRMLLG
		į .				KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
			1	1		KATSPLSLDAGLGQAPWSDLLLWALLLNRA
1	1			1		QMAMYFWEMGSNAVSSALGACLLLRVMAR
ľ		1	1.	ĺ		LEPDAEEAARRKDLAFKFEGMGVDLFGECYR
1		1	1	1		SSEVRAARLLLRRCPLWGDATCLQLAMQAD
				<del></del>	100	ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	. 403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSOREMVTGERSPSPEEEEEEEEGFGERA
1		1	1	1	1	SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
1 .		1	.	1		PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
1	1	1	ł	ţ	1	GGLRVRLP
L	1,,,,	<del>  </del>	10700	+,	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA
260	1610	A	2728	1	4//	LEYLORYHIHRDIKPDNILLDEHGHVHITDFN
				1		IATVVKGAERASSMAGTKPYMAPEVFQVYM
		1		1		DRGPGYSYPVDWWSLGITAYELLRGWRPYEI
		1	1	1	1	HSVTPIDEILNMFKVERVHYSSTWCKGMVAL
1			}	1		LRK
			1			
261	1611	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine O=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	]		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	į.			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i	[	(	peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion  RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
		ł	l		1	REPVEDPVSGNVEHILTHRREEKFEHIFGSELF
	1	Ì	l	{	1	DIFVDRRVSALAVVNECGTHPQDERLGLGW
	i	1				GLGEPGSEERLFPAAITSR
		<u> </u>	<u> </u>		431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
262	1612	Α	2733	3	431	GRLVKLSLANNNLVGVHEDAFETLESLQVLE
					1	LNDNNLRSLSVAALAALPALRSLRLDGNPWL
	1	1			j	CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
	1	1	1	]	}	ESRRISLRACRRPASRV
	1	<del> </del>	1200	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF
263	1613	Α	2736	2	343	LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
	į	1	1		i	ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
		1	1			RLISPLVNLPQSPGGLEFQYQAT
	1614	A	2738	2	245	RAMLKCLREGOPPPSYNWTRLDGPLPSGVRV
264	1014	Ι Δ	2/36	2	243	DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
	ł	ì	1	ł	ì	DTVDVLDPPEDSGKQVDL
265	1615	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
203	1013	) A	2132	12	300	LAAVETTVLVLIFAVSLLGNVCALVLVARRR
	1	1				RRGATACLVLNLFCADLLFISAIPLVLAVRWT
	1				ì	EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
		1	1	1		SLER
266	1616	A	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
200	1010	1	2133	172	1	LOGYLPLQDAFHIFQDPLTGDLPWPELILGLP
ı İ	1		İ			V
267	1617	A	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
207	1017	1.	2,00	1	1	HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
			}			LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	A	2762	1	405	IACTFCGQDEWSPERSTRCFRRRSRFLAWGEP
200	1010	1.	1 -70-	1 -		AVLLLLLLSLALGLVLAALGLFVHHRDSPL
	į	İ	1	}		VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
ĺ		1				ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
	ł	-				LPLSWAE
269	1619	A	2772	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL
] =0,	1	1				LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
	1	l	}	1		LSKNLSFSEFCFDVSY
270	1620	A	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
	1	1	}	l	.1	VEQIAKAEETHSSLSQELQARLQTVTREKEEL
}	1	1		1		LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
ļ	i		]	j	}	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
	İ	ļ	į	i		KAYDELRLQSEAFKKHSLDLLSKERELNGKL
						RHLSP
271	1621	A	2795	1	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
ſ	{				1	FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
}	1	1	ļ	1 -		RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
1	1	1			l	FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
	l			1	1	VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE
	L					SQRDDGKGALNIVEFLRKREELHQILSTVKQP
272	1622	Α	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
		}	1	1	1	RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
				1	j	GYRKVVSNNCTDGVREQYTAKPQKCPGKAP
1			i	1	j	RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
	1		,	1		VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
	_L					YQNXGIXRXTVQVDNSLGS
273	1623	A	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
1	1			1	(	DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
1	.		1	1	1	KADSLNVSRNSVMQELSELEKQIQVIRQELQL
L						AVSRKTELEEYH
274	1624	. A	2805	168	320	ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE
						IFIARNGVVGETLTHCKRV

OFO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1,00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иепсе		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
202100		]		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		l	peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
275	1625	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
	_					MGKIIFQ
276	1626	A	2813	41	266	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
						KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY
	<u> </u>	J	İ	<u> </u>		QVGPVRRNGEAGPG
277	1627	A	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
		1	1	1	ļ	LFISYLHTPKHKQHEVLQAMGSILGITGEEME PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
•	}		1		Í	GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
		ļ	]		1	LPPHNSPGKIK
	1	<u> </u>		-	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
278	1628	A	2821	238	457	VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
	1		1	1		PVDONPRLV
000	1620	<del> </del> _	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
279	1629	A	2822	342	1 1	TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS
	1	ļ		i		CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
			ĺ	1		VPSORHPTXPPPAS
280	1630	A	2825	307	77	PSMVWSYHWGVKQKRLALCVFSFEEGGRRK
280	1030	A -	2023	.) 507	1 "	CGQYWPLEKDSRIRFGFLTVTNLTGAVGEPG
	,			ł		VAFOCDGORRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
201	1031	1 "	2027	10.		ODCTPSLKSLVATGNLLDLEETAKAPLSTVSA
	1	1	1			NTTNMDEVPRPQALSGSSVVWVSGCVASRS
				1		VILSLTSG
282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
202	1000	1				TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
				1		YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
		1	}	1		YLKTLPPYYL
283	1633	A	2835	462	148	VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
		1				MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
			1		1	PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE
		1				SSEESAP
284	1634	A	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
	1			Ì	Ì	DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
		ļ				KSLAETVLNFPLDKSLLLRCSNWDAETLTED
1	1					QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
					001	KR PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
285	1635	A	2843	20	271	VCDRVSEDGINRQQAQEWCIKHGFELVELSP
		- }				EELPEEDGKCLCVRRKYGTYI
	1000		2046	107	270	TAEDVLTVAYEHGVNLFDTAEVYAAGK
286	1636	A	2845	197	278 427	FVAEVRREWAKYMEVHEKASFINSELHRAM
287	1637	Α	2851	2	444	NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
1	Ì	1				AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
			1			QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
			1			KVYLEONLAAQDRVLCALT
200	1638		2859	12	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
288	1038	, A	2009		-107	LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
1		1		1		TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
Į.			1	1		KLNELLEAIKSRDLLALIQVYAEGVELMEPLL
	1	- (	1	1		EPGQELAETALHLAVRTADQTSLHLVE
289	1639	- A	2861	12	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
209	1039	1 ^	2001	1-	1 .5.	DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
	1		1	1	1	DSGLWRMHLMEGELPASMSGSCGACINGKL
	1	1	}	1		YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
	1	- [				ITDFEGOPPTPRDKLSCWVYKDRLIYFG
1000	1640	A	2868	$+_1$	378	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI
1 790		1 4 *	1 -300	1 '		SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF
290	1	- 1	ł	•	ŀ	PDCASCLQAQDPLCGWCVLQGRCTRKGQCG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	•	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
4			Į.	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		}	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l	<b>.</b>	peptide	] -	/=possible nucleotide deletion, \=possible
			1	sequence	}	nucleotide insertion
		<del> </del>	<del>                                     </del>		<del>                                     </del>	RAGOLNOWLWSYEEDSHCLHIQSLLPGHHPR
		1	į.	ļ.		QE .
291	1641	A	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
471	1041	11	) 20,0	1 *	1	PPTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
	1	ŀ	1	ļ		GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL
	1	l		1		AKVINAENAAHKSEKFRAMATRTRQEYLKD
		1		1		LA
292	1642	A	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
292	1042	Ι^	2077	"	100	PPPPAVPYSPRYVAVHCHGMLVSCWCHL
	1,242	A	2878	<del>  1</del>	427	REKEEEVEEEDKVVKETEKEAEQEKEEDSL
293	1643	A	20/0	<b>,</b> *	72/	GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF
,		j			}	LSPEKLTAENRYYCESCASLQDAEKVVELSQ
ł .		ì	1	1		GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL
		1	1	}	1	LRLPLAGGRGOAYDL
	1644	<del> </del>	2070	109	245	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI
294	1644	A	2879	109	243	IIVFVTGGVLG
	1	<del></del>	10000	-	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
295	1645	Α	2880	3	320	NNCVGEQNHRFFCALHCKSKHFCIEFTLNTNF
		1	1		1	FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS
	ŀ	1	ì	1		
	<u> </u>	<u> </u>		<u> </u>	120	LSESISQ SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
296	1646	Α	2892	209	363	
					1	RLQEFSQKMDQVRGHWPVST
297	1647	Α	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPFMAGG
			Ì	1		KLYSTMGRFLRDRKNPACREMAVVLLANLA
1			1			QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT QIQQSQASLLHMHNPPFEPTSVDMMRRACRA
				1	ļ	
					J.,,,	LLALAKVDDNHSEF FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS
298	1648	Α	2894	310	445	
					ļ	SGLLNASAQVNL KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL
299	1649	Α	2898	1	492	KIKAKNLINYDLCSIPLGISILLV W VGVIRIL
	1		1	j		GYFQAYNVLILTMQASLPKVLRFCACAGMIY
		1			Ì	LGYTFCGWIVLGPYHDKFENLNTVAECLFSL
	1	1			<b>\</b>	VNGDDMFATFAQIQQKSLVWLFSRLYLYSFI
	İ					SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD
						LQEF
300	1650	A	2901	1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST
	j	}	1	1	1	TVTVRFVNKADFPKVRAKEQTFMFPENQPVS
	1	1		1	1	SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ
1		-		Ì		LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP
}	1	)	}			FSSYEKLDITVLDVNDNAPIF
301	1651	A	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE
		1				EPCGWMYDHAKWLRTTWASSSSPNDRTFPG
		-	i		j	KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	A	2909	2	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV
	1	1		1	1	EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
1				1		CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN
1		1		1		SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
1	-	1		1		LVNRQDRAHFM
303	1653	A	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
100	.055	1 **		1		VPPTSILEHLQRRKIMKRPSSCS
304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP
304	1034	^	1	1 - //	1 '	NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	A	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
دەد	1022	1^	2230	133	1.50	DFGLFSISGVLQAGRREDKLRIQNGWLCHLA
			1			PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE
1	1	1	1			LHAREWP
L	+	A	2944	12	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
200				1 6	1 347	
306	1656	Α.	2544	-		EIGKWLSCSLLSFPSPLAVLITTFCIVTVLGREA
306	1656	A	2,544	1		EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN 09/496	nucleotide location correspondi	location corresponding to last amino	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first amino acid residue of	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
307	1657	A	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT CTAENAVGRARRVHLTILVLPVFTTLPGDRS LRLGDRLWLR
308	1658	A	2951	1	407	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM DSSLPEEEEDEDKEAINGSGNAENRERHSESS DWMKTVPSYNQTNSSMDFRNYMMRDETLEP LPKNWEMAYTDTGMIYFIDHNTKTTTWLDP RLCKKAKAPEDC
309	1659	A	2954	2.	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE LOGAVRGGAVGGSRACQRARPRGAVLG
310	1660	A	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI TGFVQLSISVTALTAILKYGQVLMHSHVVIIW LFLAVYAVATIMFCF
311	1661	A	2963	3	465	MKPOMPGLGAPNGYGPGRGRAGVPGGPERR PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK PQKPGLRGTLKPQKSGHGHENGPWPGPCNA RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ KI
312	1662	A	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ
314	1664	A	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM
315	1665	A	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT
316	1666	A	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTIYAIDLGALTKIRIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE
317	1667	Ā	2981	3	440	VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKYTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  STLALSHSAQVLASASGRSSTTAHCQIRVWD
						VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL GDHDGRTLALWGTGHL
320	1670	A	3000	693	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA PPFNQGFCSVYITLLNELDEAVQFSNASYEAA ILENLALGTEIVRVQAYSIDNLNQITYRFDAY TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF LFPWLFLQVEVIKKAYMQGEVEFEDGENGK DGAASPRNVGHNIYILAHQLARH
323	1673	A	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND ERVFGKRGF
324	1674	A	3020	523	797	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT FCLNKEALKDEFE
327	1677	A	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ EGDLVEVVLSASATFEDFQIRPHALTVHSYRA PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK RC
328	1678	A	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART LTGALNAHNKAAVDWGWQGLIAYGCHSLV VVIDSITAQTLQVLEKHKADVVKVKWAREN YHHNIGSPYCLRLASADVNGKIIVWDVAAGV AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI HPPNYIVLWNADTGTKLWKKSYADNILSFSF D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL HRMAEKVGADITVLREREVDYDSDMPRKITE VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL LGVLTQGELDNGRGRARLNLFRHLHEIQSGR TSSISFEILGFNSKGEVHGINGTQWGQTLRMG W
330	1680	A	3040	3	397	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ VVLTMTNMGPVDTATYYCAQFARGARGSN WFDPWGQ
331	1681	A	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK GENRKTLISGMIDEPHAIVVDPLRGTMYWSD WGNHPKIETAAMDGTLRETLVQDNIQWPTG LAVDYHNERLYWADAKLSVIGSIRLNGTDPI VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV FKIHKFGHSPLVNLTGGLSHASDVVLYHQHK QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN GGSCFLNARRQPKCRCQPRYTGDKCELDQC

OPA TO	CEATE I	Mark	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	под	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
	1	İ	i	sequence		nucleotide insertion
						WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC
		1				TQQVCAGYCANNSTCTVNQGNQPQCRCLPG
	]	]	ļ	ļ	1	FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
		ì .		Ì		CRCTAYFEGSRCEVNKCSRCLEGACVVNKQS
		1	1		Ì	GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT
		1	1	}	j	MNSKMMPECQCPPHMTGPRCEEHVFSQQQP
					ļ <u></u>	GHIASILIP TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG
332	1682	Α	3045	3	952	AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT
		ł	1	1	1	LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD
	1	l		ļ		MPKSPFKRRSMNEIKNLQYLPRTSEPREVLF
į	}	}	}		1	EDRTRAHADHVGQGFDWQSTAAVGVLKAV
	l	1	}	i		QFGEWSDQPRITKDVICFHAEDFTDVVQRLQ
	1	ļ		i		LDLHEPPVSQCVQWVDEAKLNQMRREGIRY
1		1		ļ		ARIQLEDNDIYFIPRNVIHQFKTVSAVCSLAW
			Ì			HIRLKOYHPVVEATQNTESNSNMDCGLTGKR
		İ	1	1		ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP
		}				PASE
333	1683	TA-	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL
333	1003	<b> </b> ^	3070	1 '''	1	YYDGCAMIAMNĠSVFAQGSQFSLDDVEVLT
	ì				1	ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG
ļ		Į.			1	CSSKTWKVAPFVRAWWRP
334	1684	I A	3053	37	276	VITDLEEOLNOLTEDNAELNNONFYLSKQLD
334	1004	1	3033	1 "		EASGANDEIVQLRSEVDHLRREITEREMQLTS
			l			QKQVRRVNKVVRSLEDF
335	1685	A	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR
1 333	1000	1		1		NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA
}			1	ł		YNDVQYQGHYYEWLPRYNDPAAPCALKCH
1		1		1	ļ	AQGQNLVVELAPKVLDGTRCNTDSLDMCISG
	1	i	1	1	1	ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC
j		1				RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI
	'	1			]	TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV
İ				1		ENTTVEFQRGSERQTFKIPGPLMADFIFKTRY
ì						TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT
1	i	1				CGGG
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA
				Ì		LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG
					ļ	AWDWRLGSPACPHWGLHKLPRLWDPLSLYP
<u></u>					1	VLCWGT - VCCV TREIST TRRES ATVEROUNCE VOOE
337	1687	Α	3059	2	709	ILTSLVELTRFETLTPRFSATVPPCWVEVQQE
					1	QQQRRHPQHLHQQHHGDAAQHTRTWKLQT
	1	1			1	DSNSWDEHVFELVLPKACMVGHVDFKFVLN
	(	-	1	1		SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF
1	1			i		LEDHKEDILCGPVWLASGLDLSGHAGMLTLT
1				ļ		SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI
1		1	1	1	1	CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK
				<del>-  </del>	1204	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK
338	1688	Α	3060	85	384	DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL
				Į.		SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE
		- 1	1		İ	EELNP
			<del> </del>		362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV
339	1689,	A	3063	236	362	PSSVTTMLSWV
	1200	<del> </del>	12000	3	1249	DLWOFTPLHEAASKNRVEVCSLLLSYGADPT
340	1690	A	3065	13	1249	LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL
	)	- 1	}	}	1	QAAREADVTRIKKHLSLEMVNFKHPQTHETA
-		'		1	1	LHCAAASPYPKRKQICELLLRKGANINEKTKE
	- [	- 1	1	-	1	FLTPLHVASEKAHNDVVEVVVKHEAKVNAL
'			1	1	1	DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN
			1	1		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IISLQGFTALQMGNENVQQLLQEGISLGNSEA
						DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE GRQSTPLHFAAGYNRVSVVEYLLQHGADVH AKDKGGLVPLHNACSYGHYEVAELLVKHGA VVNVADLWKFTPLHEAAAKGKYEICKLLLQ HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR GDAALLDAAKKGCLARVKKLSSPDNVNCRD TQGRHSTPLHLAGK
341	1691	A	3070	1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI RTVKFLRSATIPVVELMDVQGERLDMEVGFD NRQAAFDMVCTMLEKRVRHKILYLGSKDDT RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI HLRMOLMRDALSANPDLDGVFCTN
342	1692	A	3073	463	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE IDDNRVTAEEVDILLREGEKLAPVMAKTRILR AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE RDGLDGFITITGGKLMTYRLMAEWATDAVC RKLGNTRPCTTADLALPGSQEPAKVP
343	1693	A	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS LGASRAQVLWFVILPGALPEILTGLRIGLGVG WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLQSLQVAAISTFCCLLIGYPLAWAV AHSKPSTRNILLLL
345	1695	A	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC VINQAAATGGDARQLVGYLVSQSGLPLDTSA LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL DRKALPLPELKAQAPGRAPKAGSETIIAAAFS SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	A	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI ALSQSRSQKELEDVLYSNLEELTRMAKMVSD MLFLAQADNNQLIPEKKMLNLAHEVGKVFD OFEALPE
347	1697	A _	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI
348	1698	A	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV
349	1699	A	3087	2	249	GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR SALTYEPIFLIILEALRGVIHATPCRVSLSLWD GLDLP

NO: of nucl- eotide seq-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide seq-	peptide	noa		l negnining		
eotide seq-				nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-			in USSN	location	corresponding	[=Isoleucine, K=Lysine, L=Leucine,
- 1	seq- uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	derice	ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
HELICE		}	) 214	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ł	peptide	Jodgaesse	/=possible nucleotide deletion, \=possible
			ļ	sequence		nucleotide insertion
352	1702	A	3110	341	2	AQLAEVCPPQTLLTTNTSSISITAIAAEIKNPER
332	1702	^	3110	342	} ~	VAGLHFFNPAPVMKLVEVVSGLATAAEVVE
		·	i		· ·	QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
1		1	Ì			YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
222	1705	1	] 3	-	1.00	GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	A	3116	367	225	WQLFHLNGTFLNIGETDTESCVNGWVYDRSS
334	1704	Α	3110	30,	) <b></b>	FPFSNMTEVRGLVFLS
355	1705	A	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
333	1705	^	3117	101	1 33	EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
-		1			}	ESRICVVGENGAGKSTMLKLLLGDL\APVRGI
		[	1	ĺ	ì	RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
					İ	LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
1		l	1			SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
	•	İ	Ì		<b>f</b>	DEPTN/HLGHGRAIEALGPCLQTISGVGVILVS
		i		ļ	,	HE*SALSRLVCRE\LWVC*GRSTSPF
356	1706	A	3121	137	466	RGGRDWGEHNORLEEHOARAWOGAMDAG
330	1700	^	3121	13,	100	AASREHARWQGTGLAPGTRVAVAPTCVQGL
		]	ļ	) ,	)	POERSVCRPFFSSRWREGPVWALGAGAHGKP
		1	]	'	Ì	RWSGGVRCVVRGGRWFTPAPH
357	1707	A	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ
337	1707	1^	3124	1245		PGLYFGGAAAVAEPDHLREAGITAVLTVDSE
	[	1	1			EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
	1	1				LDRCVAFIGQARAEGRAVLVHCHAGVSRSV
	1		ì		1	AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN
		i			{	EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ
						KVTEKYPELQNLPQELFAVDPTTVSQGLKDE
İ		1				VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
ł	1	1				KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
	١,	٠.	]			LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
1					İ	GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	A	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
336 .	1700	'   **	3,2,	0.0	1	TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP
	ĺ	[		Ì		LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
	1	1	1	Į.		GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
-	1	1				PTANREINPGPAAAADTRSCWGHKRSWRGW
		1		1		RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
		- {				KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
		ļ				VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
	1	]			)	HRPLDKKREDAPNLRPALAD\ITVCDYRAQIA
	1					*AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT
	1	1				AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
	1 .	1				*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
		1	1	1		QA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
1	1 -/	1				AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
į.	ì	i	l		Ì	VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N
	i i	1	- 1			GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
1	1	- 1		1	1	GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
1	1	1	1			PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ
1	1	1				APGLPHRTSIRPGWRRLTEPEAWARRHRRPW
1	1	1		1		GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
			1	i	1	PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ
	1	j	ł	1		
					ļ.	SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
						SS/GALLWTPPTTPRGSHSPRPREAPLRAIHPA GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
			1.		i.	SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RNQSPLGNDTLSSGLPMGPRRQVWPLARVG GHSSPREPQVLKKPLWGQTDIAGVGSASLYP DNL  RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ
362	1712	A	3136	1270	214	HNTWQLSRVYPSDLRTDSSNYNPQELWNAG CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGESCSSFLLGPSVAMMMQTAGL EMDICDGHFRQNGGGGYVLKPDFLRDIQSSF HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNG FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW KSRNDLLGKTPCFGTCMQQGYRHIHLLSKDG ISLRPASIFVYICIQEGLEGDES
363	1713	С	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	A	3140	57	418	SAPKTLOLPAFSLYFDLGSLKLLILRIHTSIVK NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI SGASLCPVLSPPR*PLQPSSL
365	1715	A	3145	122	413	LLPYPSLFVFLRQCHFVT/RLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIF/VF LV*TGFHYVAQAGLELLTA*S/PPQLPKVVGL QA
366	1716	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLBIEVNDLRERFSAASSASKI LQERIBEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*INPRGTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	A	3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide beginning nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide location corresponding to last amino acid residue of peptide sequence  NO: of nucl- octide seq- uence  NO: of nucl- octide location corresponding to last amino acid residue of peptide sequence  NO: of nucl- octide location corresponding to last amino acid residue of peptide sequence  N=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Giutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, V=Tryptophan, Y=Tryrosine, X=Unknown, *=Stop codon, P=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion  OLQDQLRDAQQQVKALGTERTILEGHLA QAQAEQGQQELKNLRACVLELEERLSTQE VQELQKKQVELQEERRGLMSQLEEKERRI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVI LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQ VFEEIAMLVQSALDGYPVCIFAYGQTGSGG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCEI	GL LQT EER LPV RSD
NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide seq	GL LQT EER LPV RSD
nucleotide sequence    Deptide sequence   Deptide sequence   Decation corresponding to last amino acid residue of peptide sequence   Deptide seque	GL LQT EER LPV RSD
lectide sequence	GL LQT EER LPV RSD
uence uence of peptide seq	GL LQT EER LPV RSD
uence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  915  Possible nucleotide deletion, possible nucleotide insertion  QLQDQLRDAQQQVKALGTERTTLEGHLA QAQAEQGQQELKNLRACVLELEERLSTQE VQELQKKQVELQEERGLMSQLEEKERNI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVE LPGEPTPPPGLLLFPSGPGGPSDPTRLSLS ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	GL LQT EER LPV RSD
amino acid residue of peptide sequence	GL LQT EER LPV RSD
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QLQDQLRDAQQQVKALGTERTILEGHLA QAQAEQGQQELKNLRACVLELEERLSTQE VQELQKKQVELQEERRGLMSQLEEKERNI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRLHNQLQELKGNIRVFCRVF LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEIAMLVQSALDGYPVCIFAYGTGSG TMEGGPGGPDQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRIF GGGCCEIRRAGPGSEELTVTNARYVPVSCE	GL LQT EER LPV RSD
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QLQDQLRDAQQQVKALGTERTILEGHLA QAQAEQQQELKNLRACVLELEERLSTQF VQELQKKQVELQEERRGLMSQLEEKERN SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVF LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQ VFEEIAMLVQSALDGYPVCIFAYGQTGSGG TMEGGPGGPDQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRIF GGECEIRRAGPGSEELTVTNARYVPVSCE	GL LQT EER LPV RSD
peptide	GL LQT EER LPV RSD
sequence  nucleotide insertion  QLQDQLRDAQQQVKALGTERTTLEGHLA QAQAEQGQQELKNLRACVLELEERLSTQE VQELQKKQVELQEERRGLMSQLEEKERRI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVF LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	GL LQT EER LPV RSD
QLQDQLRDAQQQVKALGTERTTLEGHLA QAQAEQGQQELKNLRACVLELEERLSTQE VQELQKKQVELQEERRGLMSQLEEKERRI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVE LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQ VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRE GGECEIRRAGPGSEELTVTNARYVPVSCE	GL LQT EER LPV RSD
QAQAEQGQELKNLRACVLELEERLSTQE VQELQKKQVELQEERRGLMSQLEEKERRI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRELHNQLQELKGNIRVFCRVE LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRE GGECEIRRAGPGSEELTVTNARYVPVSCE	GL LQT EER LPV RSD
VQELQKKQVELQEERRGLMSQLEEKERRI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVF LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	.QT EER LPV RSD
VQELQKKQVELQEERRGLMSQLEEKERRI SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVF LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	.QT EER LPV RSD
SEAALSSSQAEVASLRQETVAQAALLTER LHGLEMERRRLHNQLQELKGNIRVFCRVF LPGEPTPPGLLLFPSGPSGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	EER PV RSD
LHGLEMERRRLHNQLQELKGNIRVFCRVF LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	RSD
LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS: ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	rsd
ERRGTLSGAPAPPTRHDFSFDRVFPPGSGC VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	ا حد
VFEEIAMLVQSALDGYPVCIFAYGQTGSG TMEGGPGODPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	nn l
TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	DE
TMEGGPGGDPQLEGLIPRALRHLFSVAQE QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCE	KTF
QGWTYSFVASYVEIYNETVRDLLATGTRI GGECEIRRAGPGSEELTVTNARYVPVSCEI	lsg i
GGECEIRRAGPGSEELTVTNARYVPVSCE	GO
GGECEIRRAGEGSEELIVINARIVEVSCE	CEV
The state of the s	- C
DALLHLARQNRAVARTAQNERSSRSHSVI	Ų٢
QISGEHSSRĞLQCGAPLSLVDLAGSERLDI	GL
ALGPGERERLRETQAINSSLSTLGLVIMAL	.SN
KESHVPYRNSKLTYLLQNSLGGSAKMLM	FV
NISPLEENVSESLNSLRFASKVEPSVLFGTA	202
NISPLBENVSESLINGLEFASKVEFSVEFG	יעט
NRKWKTDPDLCVCVCVCVCVCVCVCVC	VP
MSMYRVRGGRVAGGCFIGWRAPCPRAIK	
-369 1719 A 3165 365 12 GYTSQGRWIDIERGPLTANTESLHENNFN	ALP
GYIRKIE*I*IYKKN*INFGGVGLLNIVKISI	.s/ĸ
IYRFDAIPVKILTRFFINLDKLILKFVLKTK	IAK
	IUIV.
NRIKTFYIMRRKKLGDSS	
370 1720 A. 3170 393 42 GASISPSAVIDGVEGLKPMQEQEAQEAGP	CLD
*HMAPEQWVAPR\RLLFRLIFSVLHALIIA	<b>AAA</b>
QSSAEEDEDPRN*GQSSEDQAPNQNGLIV	IVH
RVHVPLGAAATVPVHRSHFPR	<del></del>
371 1721 A 3173 770 510 GNGGCGLSQIPPSHLGAFSRGSLLSRG\DP	KGP
PPHPVIFFVFVVE\QGFTVLARMVSIS*PCI	)PP
ALASQSAGITGVSHLARPQNLYF	
DYN IND A HOODOWN PICOFFOI CIDLY	P*S
372 1722 A 3180 381 76 RVLHIDOVPAHSSPOKREISQEFQLEIRH	N U
PDLAPSGCFLFLNLKNIFK\GTHFSLVDNV	~~
TVSTWLH/SQNAQFYKDRLNGWYHCLQI	CCL
OHY*AYVEK	
TO THE PROPERTY OF THE PROPERT	PLL
373 1723 A 3181 410 14101 RREVAGPEGKGLLASAHTMETPPLLLIS	GW
SAL VAAADARKI GIR KARATI GIR	12 TC
RCDGERDCPDGSDEAPEICPQSKAQRCQI	INE
HNCLGTELCVPMSRLCNGVQDCMDGSD	EGP
HCRELQGNCSRLGCQHHCVPTLDGPTCY	CNS
SFQLQADGKTCKDFDECSVYGTCSQLCT	NTD
GSFICGCVEGYLLQPDNRSCKAKNEPVD	₹PP
USTICOCY POLITICISM TO A CONTEMPORATION OF	TTA
VLLIANSQNILATYLSGAQVSTITPTSTRQ	TIA
MDFSYANETVCWVHVGDSAAQTQLKCA	
PGLKGFVDEHTINISLSLHHVEQMAIDWI	TGN
. FYFVDDIDDRIFVCNRNGDTCVTLLDLEI	YNP
KGIALDPAMGKVFFTDYGQIPKVERCDM	mG
NOIALDFAMORYFID I OQU KYERODY	ADA
QNRTKLVDSKIVFPHGITLDLVSRLVYW	
YLDYIEVVDYEGKGRQTIIQGILIEHLYGI	
ENYLYATNSDNANAQQKTSVIRVNRFNS	TEY
OVVTRVDKGGALHIYHQRRQPRVRSHA	
DOYGKPGGCSDICLLANSHKARTCRCRS	
DO TOR OUT OF THE PROPERTY OF	77 D
LGSDGKSCKKPEHELFLVYGKGRPGIIRC	IIVII
MGAKVPDEHMIPIENLMNPRALDFHAET	
YFADTTSYLIGRQKIDGTERETILKDGIHI	
	4VE
GVAVDWMGDNLYWTDDGPKKTISVAR	LEK 1VE
GVAVDWMGDNLYWTDDGPKKTISVAR AAQTRKTLIEGKMTHPRAIVVDPLNGWI	VYW VYW
GVAVDWMGDNLYWTDDGPKKTISVAR AAQTRKTLIEGKMTHPRAIVVDPLNGWI TDWEEDPKDSRRGRLERAWMDGSHRD	IVE LEK MYW FVT
GVAVDWMGDNLYWTDDGPKKTISVAR AAQTRKTLIEGKMTHPRAIVVDPLNGWI TDWEEDPKDSRRGRLERAWMDGSHRDI SKTVLWPNGLSLDIPAGRLYWVDAFYD	IVE LEK MYW FVT RIETI
GVAVDWMGDNLYWTDDGPKKTISVAR AAQTRKTLIEGKMTHPRAIVVDPLNGWI TDWEEDPKDSRRGRLERAWMDGSHRDI SKTVLWPNGLSLDIPAGRLYWVDAFYDI LLNGTDRKIVYEGPELNHAFGLCHHGNY	IVE LEK MYW FVT RIETI (LFW
GVAVDWMGDNLYWTDDGPKKTISVAR AAQTRKTLIEGKMTHPRAIVVDPLNGWI TDWEEDPKDSRRGRLERAWMDGSHRDI SKTVLWPNGLSLDIPAGRLYWVDAFYD	IVE LEK MYW FVT RIETI (LFW

SRQ ID No. of pipuld endoted provided and pr				020	10 - 32 - 4 - 4	Dundleted and	Amino acid sequence (A=Alanine C=Cysteine,
nuclecide seld- se	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
uence  914  914  914  914  914  914  914  91	NO: of	NO: of	hod	ID NO:			
unice unice unice unice place sequence    Methionine, N=Aspragine, P=Proline,	nucl-	peptide		in ,			
unner  1944  1944  1944  1944  1946  1944  1946  1944  1946  1947  1946  1946  1947		seo-		USSN	location	corresponding	
uence  1914 ng to first samino sid of peptide peptide peptide peptide peptide peptide sequence  1915 peptide peptide sequence  1916 peptide sequence  1916 peptide peptide peptide peptide sequence peptide sequence peptide sequence  1917 peptide peptide peptide sequence peptide sequence peptide sequence  1918 peptide peptide peptide sequence peptide sequence peptide sequence peptide sequence  1918 peptide peptide sequence peptide	1	-	1	09/496	correspondi	to last amino	
maino seid residue of sequence peptide sequence of peptide sequenc		00000	1			acid residue	O=Glutamine, R=Arginine, S=Serine,
residue of peptide sequence    V=Tyrosinc, X=Unkalown, *~Slap codon, /~possible nucleotide decidion, \*~possible nucleotide decidion, \*~possible nucleotide insertion    FERRWYDAOHQQVOSNKCRVNNAGCSELCL	UGITCE	ļ	ſ	/**			T=Threonine V=Valine W=Tryptophan.
peptide sequence sequ			<b>\</b>				V-Turceine V-Ilnknown *=Ston codon
nuclonida insertion  FERMYPAO(HQQVGSNRCEVNNAGCSSLCL  ATROSRQCACAEDQVLDADGVTCANPSYVP PPOCOG PGEFACANSIC (CREWKCODANDCLD NSDEAPALCHOPTCPSDEFECENNECTPINW LCDGNDTCONSEDESANTCSARICEPRO(PSC AGGCEPISWTCDLDDCGDRSDESANSCAYPT CPLTOFTCNINGECNINVACIONNOCCONNOCCONS DEAGCSIRSCSSTOFECNSGRCIPEHWTCDGD NCGDYSDETHANCTNOATREPGGCTTDEF QCRLDGLCTPLWRCDGDTDCMDSSDESANSCAYPT CPLTOFTCSVRGECNINACTDATPGGGCTTDEF QCRLDGLCTPLWRCDGDTDCMDSSDESANSCAYPT CPLTOFTSVRGECNINACTDATPGGGCTTDEF QCRLDGLCTPLWRCDGDTDCMDSSDESANS UNDCENNSDEENCSLACEPPSHFCANNTSVC LPPKLCORDNDCDGNSGEGLCLOGCSLIN GGCSHINCSVAPGGIVCSCPLGMELGPNHIT CQUOSYCARHLKCSQRCDNEFSVRCSCYFG WYLEPDGESCRSLDPFKPFHISNRHERRIDLH KGDYSVLYGRINTLADFHLSQSALAYMTDV VEDKLYRGKLLDNGALTSFEVVIOYGLATPEG LAVDWAGNITYWVSSILDQIEVAKLGTURT TLLAGDEHPRAIALDFROGLEWTDWDASLP REASANGAGRRTVARETGGGGWPWGLTV DYLEKERL WIDDASDATYSARYDGSGHMEVL REHERSHFAVIT LOGGVYVTDWATNTILA KANKWTGINVTVVQRTNTOFFDLQVYHSS QPMAPNCEANGGQOFCSILLCINYRRIVSC ACPHLMSLHKDNTTCYPEFKKPLLYARQMEIR GRUDDAPYTYNISTTYPDDNVTLJYDYDARE QRVYWSDWRTQARKAFINGTGVETVVSADL REAGLADWSSRALFWTSTDTNKKQINVA RILDGSFRNAVVGGLOPHGLVVHPLRGKLY WTDGDDISMANMOGSNRTLLTSGGGLEVID ARRSGLGKATALAMDKIK-WTSDTDTNKKQINVA RILDGSFRNAVVGGLOPHGLVVHPLRGKLY WTDGDDISMANMOGSNRTLLTSGGGCLEVID ARRSGLGKATALAMDKIK-WTSDTDTNKKQINVA RILDGSFRNAVVGGLOPHGLVVHPLRGKLY WTDGDDISMANMOGSNRTLLTSGGGCLEVID ARRSGLGKATALAMDKIK-WTNOTDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLVSYHEGIR GPLDPPMSDSALVYDYGGTSLAGUNDEYNRGIS GCTSKADGSGSV/LRNSTLWHMKVYDESSI GOLDHRGTNEVSVNDCSQLCLTYREGRORACA CARGMLAEDGASCRYAGYLLYSERTILKSI REGRONADSVPLRTGGGVQLCLTKRGRORACA CARGMLAEDGASCRYAGYLLYSERTILKSI REGRONADSVPLRTGGGQLCLTRANCSICR REGRONADFSSVPLRTGGGQCLCLTRANGSICR REGRONADFSSVPLRTGGGOQLCLTRANGSICR REGRONADFSSVPLRTGGROPETHMSNANLAFDHY RAGTSFGTPRRFFSDHFFONTAGNALAFDY RAGTSFGTPRRFFSDHFFONTAGNALAFDY RAGTSFGTPRRFFSDHFONTAGCHLCT LTIRGDRTPGLAVBRORYTTOTT RHTVOTGRCAFERSTUTTMSGDDHRAFKLLTYGHI WTDWVRRAVQRANKHVGSNMKLLRVDIPG QPMGILAVADHTNSCELSFCRINNGGCDLCL LTTIGGHVNCSCRGGRILQDDLTCRAVYSSCR AQDEFECANGCINSTLTCOOPPRICTACCOLCERCEDGTC	1	İ	1			sequence	
FEIRMYTAOHQOVGSNKCRVNNAGCSSLCL ATGSSGCACAEGQULDAGYTCLANRSYVP PPOCOPGEFACANSKCIQERWKCDGDNDCLD NSDEAPALCHQHTCSDEFFECNNECTPRW LCDGDNDCGNSEDESNATCSARTCPPNOFSC ASGRCIPISWTCDLDDDCGDRSDESSACAYFT CPPLTOFTCNIGACININWCDNDNDGGDNS DEAGGSSICSSTOFFACNSGCCPHWTCDGD NDCGDDYSDETHANCTNOATRPPGGCHTDEF QCRLDGLCPLRWRCDDTDCMSDESBESCS GWTHYCDPSVKFGCKDSARCISKAWVCDGD NCEDNDEENCESLACPPSHECANATSVC LPPDKLDGNDDCDGDSDEGSLCDQCSLNN GCSINGSVAPEGCKDSARCISKAWVCDGD NCEDNDEENCESLACPPSHECANATSVC LPPDKLDGNDDCDGDSDEGSLCDQCSLNN GCSINGSVAPEGCKDSARCISKAWVCDGD NCEDNDEENCESLACPPSHECANATSVC LPPDKLDGNDDCDGDGSDEGSLCDQCSLNN GCSINGSVAPEGCKDSARCISKAWVCDGD NCEDNSGEROFSVKCSCYEG WALENFOXAPHAGINGSCPCHMERICAL GCSINGSVAPEGCKDSARCISKAWVCDGD NCEDNSGEROFSVKCSCYGG WALENFOXAPHAGINGSCPCHMERICAL GCSINGSVAPEGCKDSARCISKAWVCDGA VALENFOXAPHAGINGSCPCARTSVC LPPDKLDGNDDCDGDGSDEGSLCDQCSLNN GCSINGSVAPEGCKSCAPHAGING COSTANTAR CO	1	1	ł	l			
ATPGSRQCACAEDQVLDADGVTCLANPSYVP PPQCQFGFACANSCQERWKCDOMOCLD NSDRAPALCHQHTCYSDRKCENNRCIPNW LDGGMDCGNSEDSANTCSARTCPPNQFSC ASGRCIPISWTCDLDDDCGDRSDESASCAYPT CPLTQFTCNNGCININWRCDDNGCDNS DEAGCSHSCSTQFKCNSGRCFEHWTCDD NDCGDYSDETHANCTNQATRPGGCHTDEF QCRLDGLGPLRWRCDGDTDCMDSSDESASCAYPT QCRLDGLGPLRWRCDGDTDCMDSSDESASCAYPT ORTHOLOGNSDESCESAGACTSARTCSARAWCDGD NDCGDNSDEENCESLACRPPSHPCANTSVC LPDRACCSHGCSDARCISKAWCDGD NDCEDNSDEENCESLACRPSHPCANTSVC LPDRACCDGNDEGDGEGEGLOQCSLNN GGCSINICSVAPGEGIVCSCPLGMELGPDNTIT CQLGSYCAKHLKCSKCDQNKFSYKCSCYBG WVLEPDGBSCRSLDPFKPPHIFSNRHEIRRIDLH KGDYSVLYPGGRTHALDPHLSGALYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG CHAWMAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG CHAWMAGNIYWVSSINLQIELAYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG CHAWMAGNIYWTDGROBLGVITNI RGHERISHPAVIT.HGGGYVTUTDVDARS QPAMPPCEANGGGCSCHLCLINYRTVSS QPAMPPCEANGGGGCSCHLCLINYRTVSS QPAMPPCEANGGGGCSCHLCLINYRTVSS QPAMPPCEANGGGGCSCHLCLINYRTVSQ ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GYDLDAPYNYSISTYPDIDNYLDYDARE QRYYMSDVRTQAIKRAFINGTGVETVVSDL AMSGQGACHALAHAMAGNIX RLDGSFRNAVQGLEQPRIGLVYHFLRGKLY, WTDOEDNSMAMMDGSNTLLLSGQKGPVGL ADPFESKLYWISSGNITTINKCNLDGSGLEWYL WTDOEDNSMAMMDGSNTLLSGQKGPVGL ADPFESKLYWISSGNITTINKCNLDGSGLEWYL GYCTSKADGSGSVULRNSTTLVMHMKYYDGE QLANGWAGNACHAMAGNIX GREGARDATAGNIYTDQGGTDVEDVARLNG GYCTAGYSLRSGQQACEGVGSFLLYSVHGGIG RVEGIAADWIAGNIYWTDQGGTDVEDVARLNG GYCTAGYSLRSGQQACEGVGSFLLYSVHGGIG RVEGIAADWIAGNIYWTDQGTDVEDVARLNG GYCTAGYSLRSGQQACEGVGSFLLYSVHTGGIG RVEGIAADWIAGNIYWTDQGTDVEDVARLNG GYCTAGYSLRSGQQACEGVGSFLLYSVHTGGIG RVEGIAADWIAGNIYWTDQGTDVEVARLNG GYCTAGYSLRSGQQACEGVGSFLLYSVHTGGIG RVEGIAADWIAGNIYWTDQGTDVEVARLNG GYCTAGAGGACACA CAHGMLAEDQASCEYAGYLLYSERTILKS LEGCOTAGGACACA CAHGMLAEDQAS		}	_		sequence		nucleotide insertion
PPQCQPGEFACANSRCIQERWKCDGDNDCLD NSDEAPALCHQHITCSDRIKCENNRCIPNRW LCDGDNDCGNSEDESNATCSARTCPPRQPSC ASGRCPISWTCDLDDCGDRISESASCAYPT CPFLTQPTCNNGRCININWRCDNDDCGDNS DEAGCSHSCSSTQFKCNSGRCIPHWTCDGD NDCGDYSDETHANCTNQATRPPGGCHTDEF QCRLDGLCPFLWRCDGDTDCMGNSSBERSCE GVTHYCDPSVKPGCKDSARCISKAWVCDGD NDCGDYSDETHANCTNQATRPFGGCHTDEF QCRLDGLCPFLWRCDGDTDCMGNSSBERSCE GVTHYCDPSVKPGCKDSARCISKAWVCDGD NDCEDNSDEENCESLACRPSHIPCANTSVC LPPDKLCDGNDDCGGGSDEGELCDQCSLNN GCGSHINCSVAPGEGVCSCPLOMBLGPDNHT CQIQSYCAKHIKCSQKCDQNKFSVKCSCSYGE WVLEPDGBSCRSLDFFFPFIFISHRHERIDLB KGDYSVLYPGLRTHALDHLSQSALYWTDV VEDKLYTGGKLTDMALDFFFPFIFISHRHERIDLB KGDYSVLYPGLRTHALDPRDGLLFWTDWASLP LAYDWHAGNIYWESNLDQIEVAKLDGTTLRT TLLAGDEFPRALDPRDGLEWTDWDASLP RIEAASNMSGAGRTVHRETGSGGWPNGLTV DYLEKRIJMGNALLDPRDGLLFWTDWDASLP RIEAASNMSGAGRTVHRETGSGGWPNGLTV DYLEKRILWDDASSDAYSARYDGSGMEWL RIEAASNMSGAGRTVHRETGSGGWPNGLTV DYLEKRILWDASSDAYSARYDGGMEWL RIEAASNMSGAGRTVHRYTGSGGWPNGLTV DYLEKRILWDASSDAYSARYDGGWPNGLTV DYLEKRILWDASSDAYSARYDGWHAD RIEAASNMSGAGRTVHRETGSGGWPNGLTV DYLEKRILWDASSDAYSARYDGWFWLYBSR QPMAPNCEAMCGGGCSFLLCLNYNRTYSC ACPHLMKLHKDNTTCYEKKRLJYARQMEIR GYDLDAPYNYIISFTVPDIDNYTVLDYDARE QRVYWSDVRTQAKRAFINGTGVTVYASDL RIEAASNMSGAGRTVHRRTTSTDNKCQNVA RLDGSFRNAVVQGLGGGGCSGLCLNYNRTYSL QRYYWSDVRTQAKRAFINGTGVTVYASDL PRAHGLAVDWYASRLWTMSYTDNKCQNVA RLDGSFRNAVVQGLGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG				T			FEIRIMYDAQHQQVGSNKCRVNNAGCSSLCL
NSDRAPALCHGHTCYSDRYKCENNRCIPNYR LDGDDDCGNSBESBANTCSARTCPPROFSC ASGRCIPISWTCDLDDDCGDRSDESASCAYTT CPLTCPTCNNGCCINISWRCDNDNDCGDDS DEAGCSHSCSTQFKCNSGRCIPEHWTCDD NDCGDDD NDCGDYSDETHAACTINQATRPFGGCHTDEF QCRLDGLCPTLRWRCDGDTDCMDSSDESCG GYTHVCDESVKEPGCEDASACISKAVCDGD NDCEDNSDEENCESLACRPSHPCANTSVC LPDXLDCDGNDDCGDSDEGELDQCSLNN GGCSHNCSVARGGIVCSCPLQMELGPDNHT CQGSYCAKHLKCSQKCDQNLSFSVCSCYBG WVLEPDGSSCRSLDPFKPRIFSNRHEIRRIDLH KGDYSVLVPGLNTHALDHHLSQSALYWTDL VEDKIYRGKLLDNGALTSPEVUJCYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLKR TLLAGDEHPRALDPROGLEPTWDDASLP REAASNSGAGRTVHRETGSGGWPNGLTV DYLEKRILWDDASDAYSAX POLSOGHMEVL RGHEFLSHPFAVTLYGGGVYWTDWRINTLA KANKWGHNYTVVGNTTOTPEDLQVYHPSR QPMAPPGCANGGQGCSHLCLNNYNRTVSS CHARLKCHANDAYSAY DOLSOGHMEVL RGHEFLSHPFAVTLYGGGGVGCHCLANYNRTVSS CACHHMKLHKDNTTCYEKKELLYJAQMEIR GVDLDAPYNYISSTVPDIDNVTVLDYDASE GVDLDAPYNYISSTVPDIDNVTVLDYDASLP RLAGADSGSVALGCEPGGLVYLPLRGKLY WTDODNISMANMDGSNRTLLFSGQKGPCUL AMPFESKLYWISSGNHTINRCH.DGSGLEVID AMRSQLGKATALAMMGSNRTLLFSGQKGPCUL AMPFESKLYWISSGNHTINRCH.DGSGLEVID AMRSQLGKATALAMMGNLWWADQVSEC GPLDPANKSDALALAMMGSNRTLLFSGQKGPCUL AMPSEKLYWISSGNHTINRCH.DGSGLEVID AMRSQLGKATALAMMGSNRTLLFSGQKGPCUL AMPSEKLYWISSGNHTINRCH.DGSGLEVID AMRSQLGKATALAMMGNLWWADQVSEC GPLDPANKSDALAVRSTTLVMFUNDS GYPTVYDGGLGSGLCLFTSETTSS CMCTGAYSRSGQOACEGNOSTLLYSVHEGIG GPLDPANKSDALAVRSTLLAWHENE GYPRIFESRLJOTTSRAACHONINGVYDGIG RVGGAYDNIGHTSVALANTYNTING SFGYVVISGCLORGACGACGNOSTLLYSVHEGIG GPLDPANKSDALAVRSTTLVMYSISWPNGIS GYDLOPANTSDALAVRSTTLVMYSISWPNGIS GYPTRYNGFRSLLOGTGRAVAVANGCCQULLYSGTTSS CMCTGAYSRSGQOACEGNOSTLAVSTLYSVHEGIG GYPRIFESRLJOTTSRAACHONINGVYDGIG RVGGAYDNGCACGACGGGGCAGGGGGACGACGACGACGACGACGACGA	1	ļ ·	1	ł	İ	1	ATPGSRQCACAEDQVLDADGVICLANPSYVP
LCDGNDCGNSEDESNATCSARTCPROJECTS ASGRCPISWTCDLDDCGDNSESASCAYPT CPFLTQPTCNNGRCININRCDNDDCGDNS DEAGCSHCSSTQFCKNSGRCIPEHWTCDGD NDCGDYSDETHANCTNQATRPROGCHIDE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESSCE QCRLDGLPTLAWRCDGDTDCMSSDESCE LOPENTSTEAM GGCSINCSVAPGEGIVCSCPLGMELGTDNIT CQLGSYCAKHLKSQKCDQNKFSYKCSCYBG WVLEPDGSSCRSLDPFKPFIIFNNHEIRERDLIA KGDYSVLYPGLRTHALDHRUSGKLDWATKUTOV VEDKTYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWJAGNIYWFSSNLQDELAWTDWDASJL RIEAASMSGAGRTVHRETOSGGWPNGLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV DYLEKRILWIDARSDAYSARVDGSGHMCLTV RGHEFLSHPFANTLYGGGYVYUTVLDYDARE QCPHLMKLHKDNTTCYEFKKFLLYARQMER QCPHLMKLHKDNTTCYEFKKFLLYARQMER QCPHLMKLHKDNTTCYEFKKFLLYARQMER QCPHLMKLHKDNTTCYEFKKFLLYARQMER QCPHLMKLHKDNTTCYEFKKFLLYARQMER QCRYWBSDVRTQAKRAFNTTDGFDVKQDIVA RLDGSFRAVVVQGLEQPHGLVVFFLRGKLY WTDGDNISMAMMGSNRTLLFSQCKGPWCL AMRSQLGKATALAMGGSNRTLLFSGCKGPVLY WTDGDNISMAMMGSNRTLLFSGCKGPVLY WTDGDNISMAMMGSNRTLLFSGCKGPVLY QCPHCHERSKLYWISSGNRTHNRCNLDGSGLEVID AMRSQLGKATALAMGGNCLURYSTTLWHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTGAYSLRSGQACEGNUSTLLYSVHGGIR GPJDPNDKSDALVPNSGTSLAVGDPHARDT TYWVDMGLSTTSRAKRQTVTWEDVYNNGIR RGSGNATDSVPLARTTUNGGDDFRATT TYWDMGLSTRAKRAPQTWREDVYTNGIR RYGGIAVDWIAGNITWTTQGFDVIEVARING SFRYVYNSQLDKPATTYGBUFFHMCNUALAFDY RGGRANACPGCQCLCLTRGRGCAAC CAHGMLAEDGASCRYAGYLTYSTRTLKSI HLSDERUNAPCGCGCGLCLTRGRGTRTVL DECQNLMFWTNWREQHISIMRAALAGANVL TLEEDBRTNNGCCCDCLCL LTHQGHVNCSCRGGRILDDCTCRAVNSSIC QCPMIAVADTNSECLSPCRINGGCQDLCL LTHQGHVNCSCRGGRILDDLTCRAVNSSDE KPSYCNSRCKKTFRCSNGKCNSIMLURCUPP		1		1		ì	PPQCQPGEFACANSRCIQERWKCDGDNDCLD
LCDGINDCGNŠEDESNATCSARTCPPNOJPSC ASGRCPISWTCIDDDCGINSESASCAYPT CPFLTQPTCNNGRCINIWRCDNDDCGIDSS DDAGGSSHSCSTGPKCNSGRCPHEWTCDGD NDCGDYSDETHANCTNQATRPPGGCHTDEF QCRLDGLEPILAWRCDGDTDCMSSBERSCE GVTHYCDPSVKPGCKDSARCISKAWVCDGD NDCGDYSDETHANCTNQATRPRGGCHTDEF QCRLDGLEPILAWRCDGDTDCMSSDERSCE GVTHYCDPSVKPGCKDSARCISKAWVCDGD NDCEDNSDERCESLACRPPSHPCANNTSVC LPPNKLDGNDDCGDGSDEGLELQCSLNN GGCSIKNCSVAPGEGIVCSCPLGMELGPDNHT CQLGSYCAKHLKSOKCDNKRPSKCSCYBG WVLEPDGESCRSLDPFKPPIIFNNHEERRIDLI KGDYSVLVPGLRTHALDHRUSGKLDVSKTYCSCYBG WVLEPDGESCRSLDPFKPPIIFNNHEERRIDLI KGDYSVLVPGLRTHALDHRUSGKLDVSKTYCSCYBG WVLEPDGESCRSLDPFKPPIIFNNHEERRIDLI KGDYSVLVPGLRTHALDHRUSGLTWTDWDASIL RGASAMSGAGRATVHRETOSGGWPNGLTV DYLEKRILWIDARSDAYSARVDGSGHMEUT DYLEKRILWIDARSDAYSARVDGSGHMEUT DYLEKRILWIDARSDAYSARVDGSGHMEUT RHEASHSPAVITLYGGGYVWTDWTNTNLA KANKWGHNVTVVQRTNTQFFDLQVYHFSS QPMAPPCEANGGQPGSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQPGSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQPGSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQPGSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQPGSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQPGSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQPGSHLCLYNYRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQCSHLCLYNYRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR QNAPPPCEANGGQCSHLCLYFICTTRSC GRIDDPINGTNYNGDCSQCLCLFTSETTRSC MCTAGYSLRSQQACGGGVSHLLYSVHGGIR GTSLSAADSGSVYLNNSTTLVWHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLFTSETTRSC MCTAGYSLRSQQACGGGVSSHLAVRYDLYNNGIS RYSGAADSBVYLNSTTLVWHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLFTSETTRSC MCTAGYSLRSQQACGGGVGSHLAVSTLYSHEGIR GYPLDPSKLYWSGARTTHRCALDSGRIV TYDWMGLSTSRAKRQTVWTGGGPDLVEVARLN SRYVYNSQLLDKPATTYNGGDTBAAC CAHMLAEDGASCRYAGYLLYSRFTLKSI HSDERNLAPVGPEDPEHMKNNALAFDYN RGGSVEGLAYHRGWDTLYWTSVTTTTT RHTYDQTRRGAFERETVTMSGDDHPRAVIT TRHTYDQTRRGAFERETVTMSGDDHPRAVIT TRHTYDQTRRGAFERETVTMSGCDDLCL LTHQGHWNCSCRGGRILDDLTCRAVNSSC ACPERCANGECHNSLITCOCYPICKDKSDE CRYSTCNSRCKKTFRQCSNGCVSNMLWC GADDDCGGGSBEPCRADGTT		ì	i	<b>\</b>			NSDEAPALCHOHTCPSDRFKCENNRCIPNRW
ASGCPISYTCOLDDOCGDRSDESASCATY CPLTOPTCHNIGECINIWRCDNDNDCGDNS DBAGCSHSCSTQFKCNSGRCPEHWTCDGD NDCGDYSDETHAACTINQATRPFGGCHTDEF QCRLDGLCPLRWRCDGDTDCMDSSDEKSC GYTHVCDESVKEGCKDSARGLSKWCGG NDCEDNSDEENCESLACRPSHPCANTSVC LPDKLDCGNDDCGSDEGELGCSLN GGCSHNCSVAREGGVCSCPLGMELGPDNIT QQGSYCAKHLKCSSCCDQWCSSYCCSCYBG WVLEPDGBSCRSLDFFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHHSQSALYWTDI WLEPDGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHHSQSALYWTDI WLEPTGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHHSQSALYWTDI WLEPTGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHHSQSALYWTDI WLEPTGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHHSQSALYWTDI WLEPTGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHSQSALYWTDI WLEPTGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHSQSALYWTDI WLEPTGBSCRSLDPFKPEIFSNRHERRID KGDYSVLYPGLRYTHALDHSQSALYWTDI KGDYSVLYPGLRYTHALDHSQSALYWTDI KGDYSVLYPGLRYTHALDHSQSALYWTDI KGDYSVLYPGLRYTHAUDASLP REASANSOAGRYTHRETGSGGWPNGLTV DYLEKRIL WDARSALYSRAY REASANSOAGRYTHRETGSGGWPNGLTV DYLEKRIL WDARSALYSRAY RGHEFLSHPFAYTLYGGFGYWTDWRTNTLA KANKWTGHNYTVQRTNTCYPFTWDASLP RGHEFLSHPFAYTLYGGGGGGCSHLCLNYNRYSCS ACPHLMCLHKDNTTCYERKRFLLYARQMEIR GYDLDAPYYNYISSTYPDIDNYTVLDYDARS QWAPNPCEANGGGGGSCSHLCLNYNRYSCS ACPHLMCLHKDNTTCYERKRFLLYARQMEIR GYDLDAPYYNYISSTYPDIDNYTVLDYDARB RAMGLGRAYALAMDGSNRTLLFSGGKGFVIL AMBSQLGKATALAMMGSNRTLLFSGGKGFVIL AMBSQLGKATALAMMGSNRTLLFSGGKGFVIL AMBSQLGKATALAMMGSNRTLLFSGGKGFVIL AMBSQLGKATALAMMGSNRTLLFSGGKGFVIL AMBSQLGKATALAMMGSNRTLLFSGGKGFVING GPLDPHDKSDALVPNSGTSLAVGIDFHAEND THYWDMGLSTIRAKRROTWREGVLFWTEW GYPRIERSRLDGTRYVLNVSISWPNGICS QLDHKGTNFCSVNNGDCSQLCLFTSETTRSC MCTGSYSRAGOACGSVGSTLAVSTTLWSHINKYYDES GPLDPHDKSDALVPNSGTSLAVGIDFHAEND THYWDMGLSTRAKRROTWREGVLFWTEW GYPRIERSRLDGTRYVLNVSISWPNGICS GPLDPHDNSDALAWNALAFDY RAGSGTTNRIFFROIHFONIQDNDGSRTLASI LSDGRAWNATHANANALAGANVL THEAGHTNATVGCREACTHYNGGCGCDLC LTGGRAWNATHACHARLYFSDATCHCOLL THEAGHTNATCGCGCDLCL LTGGGRAWNALLAKDTOP QPMGILGAVADTTRSCCHNSLTCOMYCHCKDKSGE CADDGGGGSDEFFRCDGTCC GADDGGGGSEFRCCDGTCC GADDGGGGGGFRCCDGTCCC GADDGGGGGFRCCDGTCCCCC GADDGGGGGGFRCCDGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			}		ļ	1	LCDGDNDCGNSEDESNATCSARTCPPNQFSC
CPPLTOFTCNNGRCINNWRCDMDNDCGINS DRAGCSHSCSTOFKCNGRCIPEHWTCDGD NDCGDYSDETHANCTNQATRPFGGCHTDEF QCRLDGLICTLEWRCDGDTDCAMDSSDEKSCE, GYTHYCDPSVERGCKDSAZCISKAWCDGD NDCEDNSDEENCESLACRPFANTSVC LPPDKLCDGNDDCGDGSDEGGLCDQCSLNN GCGSIRCSVAPGEGIVCSCLLGMELGPDNHT CQIQSYCAKHLKCGSKCDOPKRSWKCSCYFG WVLEPDGESCRSLDPFRPHIFSNRHERRIDLH KGDYSVLVPGLENTALDEFHLSQSALYWIDV VEDKYRGKLLDNGALTSSEVVLQYGLAFPEG LAVDWAGNIYWESNLDGIEVAKLDGTLRIT TLLAGDIEHPRAIALDFROEILFWTDWASLP RIEAASMSGAGRATVIHERFGSGWPNGLTV DYLEKRILWIDARSDAJYSARYDOSGHMEVL RGHEFLSHPFAVILVGGGVYNLDWTNTLLA KANKWTGHNYTVVQRTNTQPFDLQVYHPSR QPAMPNFCRANGQGQFCSHLMYNRTVSC ACPHLMKLHKDNTTCYFKKFLLYARQMEIR GVDLDAYYNYISSTYVEDDWTNTLLA KANKWTGHNYTVVQRTNTQPFDLQVYHPSR QPAMPNFCRANGQGQFCSHLGNYNRTVSC ACPHLMKLHKDNTTCYFKKFLLYARQMEIR GVDLDAYYNYISSTYVEDDWTNTLA RAGLOSKNAVQGLEGPHGLVVIPLRGKLY WTDODNISMANMOSNRTLLYGOKGPVGL AIDFFESKLXWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAINGDKLWWADQVSEM GTCSKADGSGSVVLRNSTILVMMKVYDES QLDHKGTNFCSVNNDGCSQLCLFTSETTRSC MCTAGYSLRSGQQACEGGGSFLLYSVHGGIR GJELDPINKSDALVPVSGTSLAVGIDFHAEDD TTYWDMGLSTISRARRDGTWREDVVTNNGIS GFRYVINSGQLKFRATIVHPERGYLFWTEW GQYFRIERSRLDGTERVVLVNNSISWMGISV DYQDGKLYWCDARTCHERDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSVSTEDDLETGENEEVV LSSNNMDMFSSSSTEDDLETGENEEVV LSSNNMDMFSSSSTEDDLETGENEEVV LSSNNMDMFSSSSTENGGGQLCL LTHGGHVACSCRGGRUQDLLTRAVSSCR AGDETCANGCGCNISLITCDGVPHCKNSSCR ADDEFECANGCGCNISLITCDGVPHCKNSSCR ADDEFECANGCGCNISLITCDGVPHCKNSSCR ADDEFECANGCGCNISLITCDGVPHCKNSSCR ADDEFECANGCGCNISLITCDGVPHCKNSSCR		1	İ	1	İ		
DEAGCHSCSTOFKCNSGRCPEHWTCDGD NDCGDYSDETHANCTNOATRPFGGCHTDEF QCRLDGLCFLRWRCDGDTDCMDSSDESKGC GVTHVCDPSVEFGCKDSARCUCDGSLNN DCCDNSDEENCESLACRPSHPCANTTSVC LPPDKLCDGNDDCGDGSDEGLOCQSLNN GCGSHCSVAPGGGIVCSCPLGMELGPDNHT CQIGSYCAKHLKCSQKCDQNRTSWKCSCYBG WVLEPGBGSKGLDPFKPHITSNRHERRIDLH KGDYSVLVPGLRNTIALDFHLSQSALYWIDV VERKYRKKLLDKOALTSEVVLQYGLATPEG LAVDWIAGNIYWYESNLDQUEVAKLDGTLRT TLLAGDEHPRAIALDPROLEFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRIL WIDASSALYSARYDOSGHMEVL RGHEFLSHPRAVTLYGGEVYWTDWRTNTLA KANKWTGHNYVVQRTNTOFPDLQVYHPSR QPMAPNCEANGQOPCSHLCLINYNTYSC ACPHLMKLHKDNTTYVERKFLLVARQMEIR QVDLDAPYTNYISFTVPDIDNVTVLDYDARE QRYVWSDVRTQAIRRAFINGTOVETVYSADL PNAHGLAVDWSRNLFWTSYTDYNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHFURGKLY WTDGDNISMAMMGSNRTLLFSGQKGPVGL ADFFESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGDGKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMHMKYDESI QLDHKGTNCSVNNODCSGLCPTSETTRSC MCTAGYSLRSQQACEGGGSFLLYSVHGIR GIPLDPNKSDALPVSGTSLAVGDFHAEN GIPLPNKSDALPVSGTSLAVGDFHAEN GIPLPNKSDALPVSGTSLAVGDFHAEN GFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYFRIERSRLDGTBEXVLVNNSISWPHGISV DYQDKLY WCDARTDKERDLETGENEEV LSSNNMDMFSSVYERDETVWSDRTAMNGSIK RGSSCDNATDSVPLETGIGVQLKDIKVFNRISV DYQDKLY WCDARTDKERDLETGENEEV LSSNNMDMFSSVFERDITVSDRTAMNGSIK RGSSCDNATDSVPLETGIGVQLKDIKVFNRISV DYQDKLY WCDARTDKERDLETGENEEV LSSNNMDMFSSVFERDITVSDRTAMNGSIK RGSSCDNATDSVPLETGIGVQLKDIKVFNRISV DYQDKLY WCDARTDKERDLETGENEEV LSSNNMDMFSSVFERDITVSDRTAMNGSIK RGSSCDNATDSVPLETGIGVQLKDIKVFNRISV LSSNNMDMFSSVFERDITVSDRTAMNGSIK RGSSCDNATDSVPLETGIGVQLKDIKVFNRIST QKGTNVCAVANGCCQQLCLYRGRGGRACA CAHGMLAEDGASCRFYAGYLLYSERTILKSI HLSDERRINAPQVPEDPHMKNVALAFDY RAGTSFTTRIFFSDHIFGNIQNDDDSRTATIT RHTTOQUTREAGASCRFYAGYLLYSERTILKSI HLSDERRINAPQVGFEDPHMKNVALAFDY RAGTSFTTRIFFSDHIFGNIQNDDDSRTATIT RHTTOQUTREAGASCRFYAGYLLYSERTILKSI LICEYDGSHRYVLLSSERVIRFFGLAVYGEHI WTDWVRRAVQRANKRINGSNMKLLRYDDPQ QPMGILAVANDTNSCELSGERRINGGCQDLCL LTHQGHVNCSCRGGRU,QDDLTCRAVNSCR AQDEGCANGCROSSTALWONG CARDCGORDSLTCHCKACGGERCXDSTAC CARGGERGERCDDGTC GARDCGGOSCERCUSDNILWCN GADDCGGOSDEFCKACAGGGCOSCANALWCN	+			1	1		
NDGGDYSDETHANCTNQATRPFGGCHTDEF QCALDGLCTLEAWRCDGDTDCAMDSSDERS.GC GYTHYCDPSVEGCKDSARCISKAWCOGD NDCEDNSDEENGESLACREPSHPCANNTSVC LPPDKLCDGNDDCGDSDEGLCDQCSLNN GGGSHNCSVAPGGGIVCSCPLGMELGPDHIT CQIQSYCARHLKCOSKCDDNKFSVKCSCYEG WVLEPDGESCRSLDPFKPHIFSNRHEIRRIDLH KGDYSVLVFGLRNTALDFHLSQSALYWIDV VERKYRGKLLDNGALTSFEVVIQYGLAFPEG LAVDWIAGNITWVESNLDJGEVAKLDGTLERI TLLAGDEHPRAIALDPROGILFWTDWDASLP RIERAASMSGAGRTVHERTGSGWPNGLTV DYLEKRILWIDARSDAJYSARYDGSGHMEVI. RGHEFLSHFAVTLYGGEVYMDWRITNTLA KANKWTGHNTVVQRTNTQPFDLQVYHPSR QPMAPPRCEANGGQPCSHLCLHYNRIVSC ACPHLMCHLKDNTTCYFEKFLLYARQMEIR GYDLDAPYNYSISTTVPDIDNYTLDYDARE QRVYWSDWTGYDAKRAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTINKKQNNA RLDGSKNAVYGGLEQPHGLVYHPLRGKLY WTDGDNISMANMGGSNRTLLFSGQKGPVGL AUPFESKLYWISSGNHTINGCLOSGLEVUD AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKAAVQGLEQPHGLOSGLEVUD AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKAAGGGSGVYLRDSGCLPTSETTRSC MCTAGYSLRSQQQACEGGGSRLLYSVHGIR GJEDPINKSDALPYGGSLLAVGUFHAEND TITWVDMGLSTISRAKRDQTWREDVYNOIG RVEGIAVDWIAGNTYTHOGPFDVEVARLNG GFRYVISGGLDKRATTVHPERGYLLYTHEW GQYPRIERSKLAVGHTPREGYTLYTHEW GQYPRIERSKLAVGHTPREGYTLYTHEW GQYPRIERSKLAVGHGFWEDVARLNG SFRYVVISGGLDKRATTVHPERGYLLYTHEW GQYPRIERSKLAVGTWREDVYNINGIR SFRYVISGGLDKRATTVHPERGYLLYTHEW GQYPRIERSKLAGTERVLVNSISWPNGIS DYQDCKLYWCDARTDKERDLETGENREVV LSSNMMMFSVSPEPDFHWRNVILAFDY RAGTSGTTRINGFFFDHEMKNVILAFDY RAGTSGTTRINGFFFDHEMKNVILAFDY RAGTSGTTRINGFFFDHEMKNVILAFDY RAGTSGTTRINGFFFDHEMKNVILAFDY RAGTSGTTRINGFFFDHEMKNVILAFDY RAGTSGTTRINGFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFFDHEMKNVILAFDY RAGTSGTTRINGFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF			ļ	j	1	1	
QCALDGLCPLAWRCDGDTDCMDSSDESKSCE GVTHVCDPSVEGGKDSAKCVCDGD NDCEDNISDEENCESLACRPSHPCANNTSVC LPPDKLCDGNDDCGDGSDEGGLZDQCSLNN GGCSHCSVAPGEGIVCSCPLGMELGPDNHT CQIGSYCAKHLKCSQKCDGMESVCSCPLGMELGPDNHT CQIGSYCAKHLKCSQKCDGMESVCSCPLGMELGPDNHT CQIGSYCAKHLKCSQKCDGMESVCSCYCEG WVLEPDGESCASLDPFKPFIFSNRHEIRRJDLH KGDYSVLYPGLRNTALDFHLSQSALYWIDV VEDKLYRGKLLUNGALTSFEVVLQYGLATPEG LAVDWAGNITWVESNIDQIEVAKLDGTLRT TLLAGDIEHPAALDPROGLIFYTDWDASLP RIEAASMGAGRRTVIRETGSGGWPNGLTV DYLEKRILWIDARSDAIYSARYDOSGHMEVL RGHEFLSIPFAVTLYGGEVYWTDWRTINTLA KANKWTGHNVTVQRTNTOPLQVYHPSR QPMAPPCEANGGQPCSHLCLINYNRTVSC ACPHLMALHSDNTTCYEFKFLLVARQMEIR GVPLDAPYYNYIBSTVPDIDNVTVLDYDARE QRVYWSDVRTQAIKRAFINGTGGFTVYSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVFIPLRGKLY WTDGDNISMANMDGSNRTLLFSQCKGPVGL AIDFFESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMHMKYYDESI QLDHKGTNCSVNNGDCSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMHMKYYDESI QLDHKGTNCSVNNGDCSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMHMKYYDESI QLDHKGTNCSVNNGDCSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGGGSVVLRNSTILVMHMKYYDESI QLDHKGTNCSVNNGDCSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGGGGVGLAVHGROFTHERD GFLDPINKSDALPVSGTSLAVGIDFHAEND TITWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWAGNTYWTDGFUVEVALING SFRYVVISQGLDKPRAITVHPEKGYLFWTBW GQYPRIERSRLDGTERVVLVNVSISWPNGISS DYQDGKLYWCDARTDKERDLETGENREVV LSSNNMDMFSVSFEDFITWSDRTHANGSIK RGSKDNATDSVPLATGIGGVLKDIKVFNRDR QKGTNVCAVANGGCQULCLYRGRGGRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERRILDGTERVLYVNTSYTTSTIT RHTTOQUTREAGASCREYAGYLLYSERTILKSI HLSDERRILDGTERFREDHTMORIQNDDGSRIT TYENVGSVGGLAYHRGWDTLYWTSYTTSTIT RHTTOQUTREAGASCREYAGYLLYSERTILKSI HLSDERRILDGTERFREDHTWSDRTHANGSIK RGSKDNATDSVFLETGIGGVLKDIKVFNRDR QKGTNVCAVANGGCQULCL LTHGGHVNCSCRGGRLQDDLTCRAVNSCR AQDEGCANGERCSTRNINGGCQDLCL LTHGGHVNCSCRGGRLQDDLTCRAVNSCR AQDEGCANGERCSTRNINGGCQDLCL LTHGGHVNCSCRGGRLQDDLTCRAVNSCR AQDEGCANGERCSTRNINGGCQDLCL LTHGGHVNCSCRGGRLQDDLTCRAVNDSCR AGDDCGGDSDEFCKNATCG	1	i	1		ļ		
GVTHVCDPSVKFGCKDSARCISSAWCDGD NDCEDNSDEENESLACREPSHPCANNTSVC LPPDKLCDGNDDCGDGSDEGELCOGLNN GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG WVLEPDGESCRSLDPFKPHTHSNRHERRIDLH KGDYSVLVPGLRNTIALDHELSQSALYWIDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWLAGNIYWVESNLDQIEVAKLDGTLKT TLLAGDEHPRAIALDPRDGILFVTUDWASLP RIEAASMSGAGRTVHRETGSGGWPNGLTV DYLEKRILWDARSDAIYSARYDGSGHMEVL RGHEFSHPFAVTLYGGGGYWYMTUTWARTNIA KANKWTGHNYTVVQRTDTQPFDLQVYHPSR QPMAPNPCEANGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLVARQMER GVDLDAPYYNYISSTVPDIDNYTVLDYDARE GVPVYSWDVTQAIKRAFINGTGVETVVSADL PNAHGLAVDWVSRALFWTSYDTNKQNVA RLIDGSFENAVVQGLEGPHGLVVHPLKGKLY WTDGDNISMASMDGSNRTLLFSGQKGPVGL ADEPSEKLYWISSGNITHCNDGSGLEVID AARSOLGKATALAMGDKLWADQVSEKM GTCSKADGSGSVLRNSTLWAHMYVDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHGGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWEEDVTNGIG RVEGLAVDWLAGNITWTPERGYLYWTE GYPHERSRLDGTERVVLVNVSISWPNGISV DVQDGKLYWCDARTLKERDLCTEENEVV LSSNNDMBWSSVSYEDFIYWSTHANGSIK RGSKDNATDSVPLRTGIGGVQLKDIKVPNRR QKGTRNCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCRETYVLVNVSISWPNGISV DVQDGKLYWCDARTLKERDLCTEENEREVV LSSNNDMBWSSVSYEDFIYWSDTHANGSIK RGSKDNATDSVPLRTGIGGVQLKDIKVPNRR QKGTRNCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCRETYNGEDPHEWNILAFDY RAGTSFGTFRKIFFSDHFFRINGFONDEGRRAFV LDECQNLMFYNRIFFGLAGVGLKDIKVPNRR QKGTRNCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCRETYNGDDDDDGSRT TVENVGSVEGLAYHRGWDTLYWSETTILKSI HLSDERNLNAPVQFFEDEHMKNVLALAFDY RAGTSFGTFRKIFFSDHFFRINGFONDEGNALAFDY RAGTSFGTFRKIFFSDHFFRANGLAUPHRAFVL DECQNLMFYTNWFEQHFSIMMAALSGANVL TLIEKDERTFRNGLAUPHRAERLYFSDATLDKIE RCEYDGSHRYVIKSEPPHFGLAVYGEHIF WTDWVRRAVQRARKHVGSNMKLLKTVIPQ QPMGILAVANDTNSCELSEPRINNGGCQLCL LTHQGHYNCSCGGGRILQDDLTCRAVNSCG AQDCEGGESBERDNKTACGCGGEFRCRNGCC CAGDCLGESPRENDGGGGLCL LTHQGHYNCSCGGGRILQDDLTCRAVNSCG AQDCEGGSBEPCNKTACGCGGFFRCRDGTC			1	i	1	l	
NDCEDNDEGGESELCOCSLNN GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT CQIQSYCAKHLKCSQKCDQNEFSYKSCSYEG WYLEPDGESCRSLDPFRFPITSNRHERRIDLH KGDYSVLVPGLRNTIALDHLSQSALYWIDV VEDKIYRGKILDNGALTSFVVJQYGLATPEG LAVDWIAGNIYWESNLDQIEVAKLDTIAT TILAGDIEHPRAIALDPROGILFWTDWDASLP RIEAASMSGAGRITVHRETGSGGWPNGLTV DYLEKRILWIDARSDAIYSAYDGSGHMEVL RGHEISHPFAVTLYGGEVYWTDWRINTLA KANKWTGHNYTVVQRTNTQFPDLQVYHPSR QPMAPNCEANGQGQPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLVARQMEIR GVDLDAPYYNISSTYPDDNYTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWYSRILFYSVDTNKKQNNA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLIFSGQKGPVGL AUBFPSSKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGSGSVLNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSTTRSC MCTAGYSLRSGQOACEGVGSFLLYSVHEGIR GIPLDPPNDKSDALVPSGTSLAVGIDFHAEND TTYVVDMGLSTSRAKROYWEDVYNDIG RVEGIAVDWIAGNIYWTDQGPDYIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLYWTED QQYPHERSRLDGTBRYVVLNVSISWPNGISV DYQDGKLYWCDARTDKERDLTFTRID RVEGIAVDWIAGNIYWTDQGPDYIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLYWTEB GSKDNATDSVPLRTGIGGVQLKDIKVPNRDR QGYPHERSRLDGTBRYVLLVNSISWPNGISV DYQDGKLYWCDARTDKERDLTCTEENNEVV LSSNNDMBFSVSYEEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGGVQLKDIKVPNRDR QKGTNVCAVANGGCQULKDIKVPNRDR QKGTNVCARAGKUNGCNGLUKNINGCCULKDIK TLEKDRIFTNGALDPTRAKLYFSDATLUKIE RCEYDGSBRYVILKSEPVHPGILAVNSCK AQDEEFGCANGCINNSLICTOCOVPHICKNKSDE KPSYCNSRCKKTRGCSNGCCVSMLUKCN		1	1	1	1	ŀ	
LPPDKLCDGNDDCGDGSDEGELCOSCLNN GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG WYLEPDGESCRSLDPFFRFIFISNHEIRRIDLH KGDYSYLVPGLRNTIALDFHLSQSALYWIDV VEDKIYRGKLLDNGALTSFEVVIQVGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TLLAGDEHPRAIALDPROGILFWTDWASLD RIEAASMSGAGRRTYHRETGSGGWPNGLTV DYLEKRIL WIDARSDAIYSARYDGSGHMEVL RGHEFISHPFAVTLYGGEVY WTDWTNTLA KANKWTGHNVTVVQRTNTQFFDLQVYHPSR QPMAPNCEANGQGCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLVARQMEIR GYDLDAPYYNYISSTYPDDNYTVLYDTARE GYDLDAPYYNYISSTYPDDNYTVLYDDNA RLDGSFKNAVQGIEQPHGLVVHPSR QYMYSDVRTQAKKRAFINGTOVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVQGIEQPHGLVVHPLRGKLY WTDGDNISMAMDGSRRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGGNLWADQVSEM GTCSKADGSGSVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSLCPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIGFFAEND TITVWVDMGLSTBRAKRQTWREDVVTNGIG RVEGIAVDWIAGNITWTDQGFDVEARLNG SFRYVISGOLDKPRAITVPBERGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTNERGVLFTETGREVV LSSNNMDMFSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVPNDRD QKTSNVCAVANGGCQLLYRGTTRSC CAHGMLAEDGASCREYAGYLLYSVRIGTI RISSNNMDMFSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVPNDRD QKGTINVCAVANGGCQLLYRGRQRACA CAHGMLAEDGASCREYAGYLLYSTPILTSTIT RITVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNNEQHFPBIMMNVIALAFDY RAGTSPGTPRRIFFSDHHFGNQLDDGSRRTI TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTSTIT TVENVGSVEGLAYHRGWDTLYWTSTYTTTTIT TTELEDGURTPNOLADHTRAKLYTSDATLDKIE RCEYDGSIRYYULKSEPVHPFGLAVNSCR AQDEFFECANGEGNELOCKSINLOCN GADDCGDGSDEIPCNXTACGCVGEFFRCNDGC	1	I	1	1	1		
GGCSHNCSVAPEGEIVCSCPLOMELEDPNHT CQQSYCARMLKCSQKCDNKFSVKCSCYEG WVLEPDGESCRSLDPFKPHIFSRIFERIDLE KGDYSVLVPGLRNTIALDFHLSQSALYWIDV VEDKYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TILAGDEIPRAIALDPRGGILFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRIL WIDARSDATYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGEVY WTDWRTNTLA KANK WTGHNVTVVQRDTYDDLQVYHPSR QPMAPNPCEANGQGPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLVARQMEIR GYDLDAYYNYNISTYPDDINVTVLDYDARE QRYVWSDVRTQAIKRAFINGTGVETVVSADL PNANGLAVDWYSRILF WTSYDTNKKQNNVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMAMMDGSRTLLFSQKGPVGL AIDFPESKLYWISSGNITTINCNLDGSGLEVID AMRSQLGKATALAMGDKL WWADQVSEKM GTCSKADGSGSVVLINSTILVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGUGFHAEND TTYWVDMGLSTIRAKRDTWREDVYTNGIG RVEGIAVDWIAGNIYWTDQGPDVIEVARLNG SFRYVVISQGLDKPRAITVHREDVTNGIG RVEGIAVDWIAGNIYWTDQGPDVIEVARLNG SFRYVVISQGLDKPRAITVHREBVGLYLPWEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSSVYEEDFYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVPNDRD QKGTNVCAVANGGCQLLYBRGGNACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLAPPQFEEDFHMNVIALAFDY RAGTSPFTRIFFSDHFGNIQQNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTYTTSTT RRITVPQTFRGAFERETVILTMSGDDHPRAFVL DECQNLMFWTNNWEQHPSIMRAALSGANVL TLEKEDRETPROLAPPGRAFFERTULNSE RCEYDGSHRYVILKSEPVHPFGLAVNGEH WTDWVRRAVGRAFKHVGSMKKLLRVDDPQ QPMGIIAVANDTNSCELSPCRINNGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGCINISLITCDGVPHCKNKSDE KPSYCNSRCKKTFRQCSNGRCVSMLWCN			1	1	ļ	ļ	
GGCSHNCSVAPEGEIVCSCPLOMELEDPNHT CQQSYCARMLKCSQKCDNKFSVKCSCYEG WVLEPDGESCRSLDPFKPHIFSRIFERIDLE KGDYSVLVPGLRNTIALDFHLSQSALYWIDV VEDKYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TILAGDEIPRAIALDPRGGILFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRIL WIDARSDATYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGEVY WTDWRTNTLA KANK WTGHNVTVVQRDTYDDLQVYHPSR QPMAPNPCEANGQGPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLVARQMEIR GYDLDAYYNYNISTYPDDINVTVLDYDARE QRYVWSDVRTQAIKRAFINGTGVETVVSADL PNANGLAVDWYSRILF WTSYDTNKKQNNVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMAMMDGSRTLLFSQKGPVGL AIDFPESKLYWISSGNITTINCNLDGSGLEVID AMRSQLGKATALAMGDKL WWADQVSEKM GTCSKADGSGSVVLINSTILVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGUGFHAEND TTYWVDMGLSTIRAKRDTWREDVYTNGIG RVEGIAVDWIAGNIYWTDQGPDVIEVARLNG SFRYVVISQGLDKPRAITVHREDVTNGIG RVEGIAVDWIAGNIYWTDQGPDVIEVARLNG SFRYVVISQGLDKPRAITVHREBVGLYLPWEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSSVYEEDFYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVPNDRD QKGTNVCAVANGGCQLLYBRGGNACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLAPPQFEEDFHMNVIALAFDY RAGTSPFTRIFFSDHFGNIQQNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTYTTSTT RRITVPQTFRGAFERETVILTMSGDDHPRAFVL DECQNLMFWTNNWEQHPSIMRAALSGANVL TLEKEDRETPROLAPPGRAFFERTULNSE RCEYDGSHRYVILKSEPVHPFGLAVNGEH WTDWVRRAVGRAFKHVGSMKKLLRVDDPQ QPMGIIAVANDTNSCELSPCRINNGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGCINISLITCDGVPHCKNKSDE KPSYCNSRCKKTFRQCSNGRCVSMLWCN	1	1	1		1	1	LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
WYLEPDGESCRSLDPFKPFIFINSHERRIDLH KGDYSVLVPGIRNTALDHISQSALYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWESNILDQIEVAKLDGTLET.  TILLAGDIEHFRAIALDFROGILFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRIL WIDARSDAYSARYDGSGMPKUL RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA KANKWTGHNYTVVQRTNTQPFDLQYHPSR QPMAPPPCEANGGQOFCSHLCLNYNRITVSC ACPHLMKLHKDMTTCYEFKKFLLYARQMEIR GVDLDAYYNYISFTDDINVYLUDYDARE QRVYWSDVRTQAKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFRAVVQGLEQPHGLVVHPLRCKLY WTDGDNISMANMGSSRTLLFSGQKGPVGL AMPPESKLYWISSGAHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLYMHMRYVDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGGVSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDPHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYVDGFDVIEVARING SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERDDLETGENREVV LSSNNMDMFSVSYEDFIFWSDRTHANGSIK RGSKDNATDSYPLRTGIGVQLKDIKVPNDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGFLLYSETILKSI HLSDERRINAPVQPFEPPEHMKNVALAFDY RAGTSFGTPNRIFFSDHFGNIGQNIDGSRRTI VENGSVEGLAYHRGWDTLYWSTYTTSTIT RHTVDQTRPGAFERSTVITMSGDDHPRAFVL DECQNLMFWTNWREQHPSIMRAALSGANVL TLEKDIRTPNGALDHAREST TIENGTSVEGLAYHRGWDTLYWSTYTTSTIT RHTVDQTRPGAFERSTVITMSGDDHPRAFVL DECQNLMFWTNWREQHPSIMRAALSGANVL TLIEKGIRTPNGALDHAREST TORMSCRAVPLIFF WTDWYRRAVQRANKHYGSNMKLLRVDIPQ QPMGILAVANDTINSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGECINFSLTCDGVPHCKNKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGGOGSDEIPCNKTACGVCFRKROGTC	+	1	i	1		1	
W\u00c4\u00e		1	1	1	1	1	COIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
KGDYSYLVPGIRNTIALDFHLSQSALYWTDV VEDKYYRKLLDNOALTAPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TILLAGDEHFRAIALDFROGILFWTDWDASLP RIEAASNASGARRTYHRETGSGWPNGLTV DYLEKRILWIDARSDAIYSARYDGSGHMEVL RGHEFLSHPAVILYGGGYYWTDWRTNILA KANKWTOHNYTVVQRTNTQFFDLQVYHPSR QPMAPNPCEANGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYFEKKFLLVARQMEIR GVDLDAPYYNYIISFTVPDIDNYTLDYDARE GRYVWSDVRTQAIKRAFNGTGVETVYSADL PNAHGLAVDWVSRNLFWTSYDTNKKQNVA RLDSFSKNAVVQGLEPGILVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLHKGTNPCSVNNGDCSQLCLFTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPYSTANTOHDAINDTHAEND TITYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWLAGNIYWTDQGFDVEVARLNG SFRYVVISQCLDKPRAITVHPEKGYLLFWTEW GQYFRIESRLDGTERVVLNVISSWPNGISV DYQDGKLYWCDARTATKHERDLETGENREVV LSSNNADMFSVSYEDFIYWSBRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRQQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNINAPVQPFEPPEHMKNVLALAFDY RAGTSPGTPRIRFSDIHFGNIQQNDDGSRRT TVENVGSVEGLAYFRGFPPEHMKNVLALAFDY RAGTSPGTPRIRFSDIHFGNIQQNDDGSRRT TVENVGSVEGLAYFRGFPPEHMKNVLALAFDY RAGTSPGTPRIRFSDIHFGNIQQNDDGSRRT TVENVGSVEGLAYFRGFPPEHMKNVLALAFDY RAGTSPGTPRIRFSDIHFGNIQQNDDGSRRT TVENVGSVEGLAYFRGFRIKKVILAKFDY CAGNINGVORDHARAEKLYFSDATLUKIE RCEYDGSHRYVLKSEPVHFFGLAYGEHF WTDWVRRAVQRANGTHARAELSGANVL TILEKDIRFTNGLANDHARAEKLYFSDATLUKIE RCEYDGSHRYVLKSEPVHFGNAVCHKINKPDIPQ QPMGIIAVANDTINSGELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGECINESLCTDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGGOGSDEIPCNKTACGVCFFRCROTC				1	1	1	
VEDKIYRGKILDNGALTSFEVVIGYGLATPEG LAVDWAGNIYWESLDQEWAKLDGTLKT TLLAGDEHPRAIALDPRDGILFWTDWDASLP RIEAASMSGAGRITVHRETIGSGWPNGLITV DYLEKRILWIDARSDAIYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA KANKWTGHNIYVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGGQPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNYIISTYPDIDNYTVLDYDARE QRYYWSDVRTQALKRAFINGTGYETVVSADL PPAHGLAVDWVSRIF,WTSYDTNKQINVA RLDGSFKNAVVQGLEQPHGLVYEPLRGKLY WTDGDNISMANMDGSNRTLLFSQQKGPVGL AMBPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGGKLWADQVSEKM GTCSKADGSGSVVLRNSTTLVMINKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQACEGVGSFLLYSVHEGIR GPLDPNKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWIAGNIYTDQGFDVEVARLNG SFRYVVISQGLDKPRAITVIFPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDAATDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLETGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFFEDPEHMKNVIALAFDY RAGTSGTTPNRIFTSDHFGNIQQINDDGSRRIT VENVGSVEGLAYWGSVFTLYSTYTTITI RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLADHRARKLYFSDATLAKIE RCEYDGSHRYVILKSEPVHFFGLAVYGEHFF WTDWVRRAVQRAKHVFSDMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLURCN	1	1	1	1		1	KGDYSVLVPGLRNTIALDFHLSOSALYWTDV
LAVDWIAGNIYWESNLOGUEVAKLDGTLRT TILAGDEHPRAIADPROGILFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRILWIDARSDAIYSARYDGSGHMEVL RGHEFI.SHFAVIT.YGGEVYWTDWRTNTLA KANKWTGHNVTVVQRTNTQPFDLQVYHPSR QPMAPPCEANGGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYSFKKPLLYARQMEIR GVDLDAPYTYNISTTVPDIDNYTVLDYDARE GRYVWSDWTROLKRAFINGTGVETVVSADL PNAHGLAVDWYSRILFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDPFESKLYWISSGNITINRCNLDGSGLEVID AMRSQLGKATALAMGDKLWWADDVSEKM GTCSKADGSGSVVLRNSTTLYMHMKYYDESI QLDHHGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQQACEGVGSFLLYSVHEGIR GPILDPNDKSDALVPVSGTSLAVGDFHAEND TTYWVDMGLSTSRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISGGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERDLETGENREVV LSSNNMDMFSVSYEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRQGRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFSGIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHRAFVL DECQNLMFWTNWNEQHFISIMRAALSGANVL TLEKDRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHFFGLAVYGEHF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSFCRNNGGCODLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGECNISLTCDGVPHCKDKSDE KPSYCNSRRCKKTRQCSNGRCVSNMLURCN	1	1	1	ì		1	VEDKIYRGKLLDNGALTSFEVVIOYGLATPEG
TILAGDEHPRAIALDPROGILFWTDWDASLP RIEAASMSGAGRTVYRETGSGGWPNOLTV DYLEKRILWIDARSDAYSARYOGSHMEVL RGHEFLSHFAVTLYGGEVYWTDWRTNTLA KANKWTGHNYTVYQRTNYQFDQVYHDSR QPMAPNFCEANGGQPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYFFKKFLLYARQMEIR GVDLDAPYYNYIISTVYPDIDNYTVLDYDARE QRYYWSDWRTQAIKRAFINGTGYETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKQDNVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AMPFESKLYWISSGNHTINRCNLOGGGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHNKVYDESI QLDHKGTNPCSVNNGDCSQLCLFTSETTRSC MCTAGYSLRSGQQACEGYGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGUFHAEND TIYWVDMGLSTISRAKRQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQFDVIEVARLNG SFRYVVISGGLDKPAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNAMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGQQLCLYRGRQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFFEDPHKNNVIALAFDY RAGTSPGTPNRIFPSDHFGNIQQINDDGSRRTI VENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERETVITMSGDDHPRAFVL DECONLMFWTNWNEQHFISIMRAALSGANVL TLEEDIRTPNGLADDHRABELYFSDATLDKIE RCEYDGSHRYVLKSEVHPFGLAVYGEHF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSFCRNNGGCQDLCL LTHOGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGECNINSLTCDGVPHCKDKSDE KPSYCNSRRCKKTRQCSNGRCVSNMLWCN GADDGGGGSBEPCRKTAGGGGFRCROSNLW.	ĺ	1	1	1:	1		I AVDWIAGNIYWVESNI DOIEVAKLDGTLRT
RIEAASMSGAGRRTVHRETGSGGWPKOLTV DYLEKRIL WIDASDATYSAR YDGSGHMEVL RGHEFLSHPFAVTLYGGEVY WTDWRTNTLA KANKWTGHNYTVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNYIISTVPDIDNVTVLDYDARE QRYYWSDVRTQAKRAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVGLEQHGLEVHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFFESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESJ QLDHKGTNPCSNMGDCSQLCJPTSETTRSC MCTAGYSLRSQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVUDEPHAEND TTYWDMGLSTISRAKRQDTWREDVYTNIGIS RVEGJAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNMDMPSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQUCLYRGRORACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFTSDHEGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRRGAFERTVITMSGDDHFRAFVL DECQNLMFWINGCHPSIRMALASGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVLKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHYNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRCKKITFRQCSNGRCVSNMLWCN	ì			1	1	1	TI LAGDIEHERAIAL DERDGILEWTDWDASLE
DYLEKRIL WIDARSDATYSARYDOSGIMEVL RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA KANKWTGHNVTVQRTNTQPFDLQVYHPSR QPMAPNPCEANGGQGPCSHLCL.NYNRTYSC ACPHLMKLHKUNTTCVEFKKFLLYARQMEIR GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE QRYYWSDVRTQAKRAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVIPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVCL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVURNSTILVHIMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPSGTSLAVAGDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGLAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDRPAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKEIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFFEDIFHMKNVIALAFDY RAGTSGGTINAPVQFFEDIFHMKNVIALAFDY RAGTSGGTINAPVQFFEDIFHMKNVIALAFDY RAGTSGGTINAPVQFFEDIFHMKNVIALAFDY RAGTSGGTINAPVQFFEDIFHMKNVIALAFDY RAGTSGGTINAPVQFFEDIFHMRAVLAFDTI RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTINNECQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHYNCSCRGGRILQDDLTCRAVNSSC AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN		1	1	1	1		DIEAA SMSGA GDDTVIDETGSGGWPNGI TV
RGHEFLSHPAVTLYGGEVYWTDWRTNTLA KANK WTGHNYTVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGQQPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYPEKFLLYARQMEIR GVDLDAPYYNJISFTVPDIDNYTVLDYDARE QRYYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSVDTINKKQINVA RLDGSKRNAVVGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFFESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDESQLCLPTSETTRSC MCTAGYSLRSGQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGANTYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPOFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVLKSEPWHPFGLAVYGEHF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTINSCELSPCRINNGGCQDLCL LTHQGHYNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCUSNMLWCN		i	İ	1	1		DATE END THE PER VICE BY DESCRIPTION OF THE PER VICE BY DESCRI
KANKWTGHNYTVVQRTNTOPFDLQVYHPSR QPMAPNPCEANGGQGPCSHLCLINYNRTVSC ACPHLMKLIKONTTCYEFKKFLLYARQMEIR GVDLDAPYYNYIISTTVPDIDNVTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMAMMGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGBKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQQACEGVGSFLLYSVHEGIR GIPLDPIDKSDALVPVSGTSLAVGIDFHAEND TITYWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVUSQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTBRVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGERNEVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLIKTIGGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCRFYAGYLLYSERTLKSI HLSDERNLNAPVQPFEDPEHMKNVLALAFDY RAGTSPGTPNRIFTSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWINWEQHEBIMRAALSGANVL TLEKDIRTPNGLADHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHEPGLAVYGEHF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNGSCEGGRILODDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KFSYCNSRRCKKITRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		i .	}	}	1	ļ	DILEKKILWIDAKSDAITSAKIDUSUNIEVL
QPMAPPPERANGGÓPCSHLCLINYNRTVSC ACPHLMKLHKDNITCYEFKKFLLYARQMEIR GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TITYWVDMGLSTRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYYVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSITSARKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSITGATRAVLNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVEEDFIYWSDRTHANGSIK RGSKDNAITDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNI.NAPVQPFEDPEHMKNVALAKDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RRITTVDQTRPGAFERETVITNSGDDHFRAFVL DECQNLMFWINWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRUGTC	Į.	1					
ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNIISFTVPDIDNVTVLDYDARE ORVYWSDVRTQAKRAFINGTGVETVVSADL PNAHGLAVDWSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQELQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGYVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GPLDPNDSADLVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISOGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNINAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDGTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHFSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSIRTYVLLKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGCQDLCL LTTIQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1	· l	1	1		ļ	
GVDLDAPYYNYISTVPDIDNYTVLDYDARE QRVYWSDVRTQAIKAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMGSNRTLLFSGQKGPVGL AMPRESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGYGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAYDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQOLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRRGAFERETVITMSGDDHPRAFVL DECQNLMFWINWNEQHFSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSIRFYVILKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGCCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECNIPSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1	1		1			
QRVYWSDVRTQAKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL ADFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVSFLLYSHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVEVARLNG SFRYVVISQGLDKRRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFTSDIH-GNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLLWCN GADDCGDGSDEIPCNKTACGVOEFRCRDGTC	ļ	1	1	ſ	i	ļ	
PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVYLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFFEDPEHMKNVLALAFDY RAGTSPFTNRIFSDIHFGNQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWINNNEQHPSIMRAALSGANVL TLEKDRITNGLADHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGGSSDEIPCNKTACGVGEFRCRDGTC	1	1	1	l.	Ţ	i	GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVYLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFFEDPEHMKNVLALAFDY RAGTSPFTNRIFSDIHFGNQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWINNNEQHPSIMRAALSGANVL TLEKDRITNGLADHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGGSSDEIPCNKTACGVGEFRCRDGTC			1	1	Į.	1	QRVYWSDVRTQAIKRAFINGTGVETVVSADL
RLDGSFKNAVVQGLEQPHGLVVIPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL ADFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNTWTDQGFDVIEVARLNG SFRYVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLEKGNIPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECNFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1	ı	1		[		PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
WTDGDNISMANMDGSNRTLLFSGGKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYYVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVLALAFDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVTIMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVLKSEPVHPFGLAVYGEHIF WTDWVRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	-	1	1	İ			
AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKATGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTE W GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYYLLYSERTILKSI HLSDERNLNAPVQFFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1			}		1	WTDGDNISMANMDGSNRTLLFSGOKGPVGL
AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNI.NAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSVTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	}		1	1		[	
GTCSKADGSGSVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RRHTVDQTRPGAFERETVITMSGDDHFRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGGVGFRCRDGTC				1			
QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQACEGYGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRABKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		1		ì	1		
MCTAGYSLRSGQACEGVGSFLLYSVHEGIR GPILDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	}	1 .	ì	ł	ł		OF DIESCENDING COOL OF DESERTE SC
GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNILNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNISSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		1		1	ŀ		CENTRALING CONTROL OF CACCACACA I ACTUAL CACACACACACACACACACACACACACACACACACAC
TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC				1	1		
RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNML.WCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	}		-	1	}	1	
SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRQGRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1 .	1		-	1		
GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			1				
GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		)	}	1	1	1	
LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			1				GQYPRIERSRLDGTERVVLVNVSISWPNGISV
LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		1	1	ſ			DYQDGKLYWCDARTDKIERIDLETGENREVV
RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTINSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			1	}	1	ł	
QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTINSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	(	1		[			
CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTINSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					1	1	OKGTNVCAVANGGCOOLCLYRGRGORACA
HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					1	1	
RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					I	1	
IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	I	1	l		1	1	
RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	İ	-	1		1	1	
DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		}		-		1	IVENVGSVEGLAYHKGWDILYWISYITSIII
TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		1	- 1	-		1 .	
RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		i	- 1		1	· l	
WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		1		}	-	1	
WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					1		RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			Ì			1	
LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	}		1	}			
AQDEFECANGECINFSLTCDGVPHCKDKSDE  KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1			1	- (		
KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1		- 1.		1		
GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1		J.		1	1	
		i	- [		1		
IGNSSKCRQF VDCEDASDEWINGSATDCSSTF			1	1	l	1	
							IGNODICHÁL ADCEDIADDEMICONIDOSO II.

			690	B 10 1	B - 12-4-11	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-`	peptide		in l	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
					to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi		
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	i	i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i		j l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	]				bodaense	/=possible nucleotide deletion, \=possible
l				peptide		
l	1			sequence		nucleotide insertion
						RLGVKGVLFQPCERTSLCYAPSWVCDGAND
		1			· .	CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
j				<u> </u>		MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
Ì	ſ	i	ł		Į	
ł	}	ł i	ł	l	i	CONHRCISKOWLCDGSDDCGDGSDEAAHCE
Į.	ì		1			GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
	ł		ł	į.		DGADESIAAGCLYNSTCDDREFMCQNRQCIP
	1		l	1	i	KHFVCDHDRDCADGSDESPECEYPTCGPSEF
		1				
1	1	1	l			RCANGRCLSSRQWECDGENDCHDQSDEAPK
ļ	1	l .	ŀ	ł	1	NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
l	i	1	l .	l	1	GQDDCGDSSDERGCHINECLSRKLSGCSQDC
1	l	1	l	1		EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
1		1	1			
	1	1	1	1	1	TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
1	1	1	,	ŀ	1	HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
1	]	}		1	1	TLLKQGLNNAVALDFDYREQMIYWTDVTTQ
i	1	1	ļ	1	1	GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
1	ł	i	1	i		
1	I	1	1	1	l	DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL
	1	l	1	i	i	VSSGLREPRALVVDVQNGYLYWTDWGDHSL
1	l	1	1	1	1	IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE
		,	1 .	1	1	RIYWADAREDYIEFASLDGSNRHVVLSQDIPH
		1	1		-	IFALTLFEDYVYWTDWETKSINRAHKTTGTN
j	}	)	] .	}		
1	1	1	1			KTLLISTLHRPMDLHVFHALRQPDVPNHPCK
1		1				VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD
1	1		1	1	1	GRTCVSNCTASQFVCKNDKCIPFWWKCDTE
	ĺ	1	i	1	i	DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
1	1		1	1	}	
	İ			1	j	PAFICDGDNDCQDNSDEANCDIHVCLPSQFK
	[	1	i	1		CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE
i	Ĭ		ł		1	VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
			1			GSDEPANCTQMTCGVDEFRCKDSGRCIPARW
	1	1	1	1		
1	ł		1			KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
1	ı	1	i		1	KNNRCVPGRWQCDYDNDCGDNSDEESCTPR
1	1	1	Ī	1		PCSESEFSCANGRCIAGRWKCDGDHDCADGS
	i	1	1			DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA
1			1	· .		DADCMDGSDEEACGTGVRTCPLDEFQCNNT
1		1 '	· <b> </b>	ŀ		
į		1	1	1.		LCKPLAWKCDGEDDCGDNSDENPEECARFV
1	i	1	1	1		CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD
1	1		1	I		GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
1	1	1	1	ł		SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT
1	1		1 .	1		NASICGDEARCVRTEKAAYCACRSGFHTVPG
1	1	1				
	1		1			QPGCQDINECLRFGTCSQLCNNTKGGHLCSC
1	l	1	1	1	1	ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS
	İ	1	1			LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA
1	1	1	1	1	ļ	GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR
1	Ì	ı	1			
1	1	1	1			RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY
1	1	Ī	1	1	ŀ	WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
1	1	1	1	1	1	AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
1	1		1	1		LRETLYQDNIQWPTGLAVDYHNERLYWADA
1	ì	1		1	1	
1	1	1	1	1	1	KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
1	1		1	1		FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG
1		1	1		1	LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
Į	1	1		1	l	CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP
		1		1	1	PDAPRPGTCNLOCFNGGSCFLNARROPKCRC
1	1		1	1	1	
I	1	1	ł	ì	I	QPRYTGDKCELDQCWEHCRNGGTCAASPSG
1	l l	1	Ī	1	1	MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT
ł	1		I		1	VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
1	1		ĺ	ł	1	NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
1	ı			I	1	
1	1		1	ł		CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
j	1	1	1	)	1	CLTCVGHCSNGGSCTMNSKMMPECQCPPHM
	Ì	1 '	1	ł		TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL
1		- 1	1 .	i	Į.	VAGVVFWYKRRVQGAKGFQHQRMTNGAM
	. ]	- [		1	ļ	
1	1	L	L			NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

NO. of No. of Load D.NO. inclosing mucle-decide social control of the control of	5-16 VB	65675	17.4	LOTO.	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
muclecide corticle soquence wence where the contract of the co	SEQ ID	SEQ ID	Met	SEQ			De Americ Acid Fe-Glutamic Acid
unice    Sequence   Se			пос				F=Phenylalanine G=Glycine, H=Histidine.
1724   1724   A   3187   191   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1724   A   3187   191   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1724   A   3187   191   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1724   A   3187   191   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1724   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1724   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSLLAPAPIMISLMPO   ATLIEUTE   A   1815   CLELASAGKIPESKALSKALSKALSKALSKALSKALSKALSKALSKALSKAL			1				I=Isoleucine K=Lysine L=Leucine.
1724			l				M=Methionine N=Asparagine P=Proline.
amino acid cecidence of peptide sequence per peptide sequence of p	-	uence	İ				
	uence			914			T-Threanine V=Veline W=Tryptophan
peptide			,				V-Typesine X=Linknown +=Ston coden
			1	İ		Sequence	/
1724   A   3187   191   1815   CIELASAGKIPESKALSLLAPATMISLMFQ   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPITPNIPTLITQUISSIGAIPAALDPN   AGILIPATTA   AGISTIAALPI OVA AGILIPAT   AGISTIAALPI OVA AGILIPAT   AGISTIAALPI OVA AGIST		1	1	ł			
					sequence		
1724   A   3187   191   1815   C.E.L.ASAGKIPEESKALSLLAPATIMISLING AGLIPTIPNELTICUSLISSLE ADAPAALDING ATLGEIPOPPLITICUSLISSLE ADAPAALDING ATLGEIPOPPLITICUSLISSLE ADAPAALDING ATLGEIPOPPLITICUSLISSLE ADAPAALDING ATLGEIPOPPLITICUSLISSLE ADAPAALDING ATLGEIPOPPLINGVISSLE AND AGABLE WAR ALGEIPOPPLINGVISSLE AND AGABLE WAR ALGEIPOPPLINGVISSLE AND AGABLE WAR ALGEIPOPPLINGVISSLE AND AGABLE WAR ARREST AND AGABLE WAR ARREST AGABLE WA		i	l				1
AGILPPTPNLTTLGVSLSSLQAPAAALDT    ATLGEPQPEMONPSKIDERRITVYYGNL     ATLGEPQPEMONPSKIDERRITVYYGNL     NSQTTADQLLEFKQVGEVKFVRMAGDEP     QFTRFAFVEFADQNSVPRALANFVMAFGDRP     QFTRFAFVEFADQNSVPRALANFVMAFGDRP     LKINHISNNAIMVKPPEMTPQAAAKELEEVMKV     VREAGSFISAALPEGWLHSTSLCADFLGGFAR     KVFPPPLVWEGGVKRRVRSSHSKDFLGGFAR     KVFPPPLVWEGGVKRRVRSSHSKDFLGSKSSSSK     SFTRKRSQSKHRSSSHNSKSSRSKDTREKTSSSSSK     SPHKKRSKSRRKKSRSSRSHKDTREKT     KKEKE VEKDREKERERKERKERKERKERKERK     KKEKE VEKDREKERERKERKERKERKERKERK     KKEKE VEKDREKERERKERKERKERKERK     KKEKE VEKDREKERERKERKERKERKERK     KKEKE VEKDREKER REKERKERKERKERK     KKEKE VEKDREKER REKERKERKERKERK     KKEKE VEKDREKER REKERKERKERKERK     KKEKE VEKDREKER REKERKERKERK     KKEKE VEKDREKER REKERKERKERK     KKEKE VEKDREKER REKERKERKERK     KKEKE VEKDREKER REKERKERK     KSTEPKSYNASRRSRSSSRERRRRRSSSSS     PTSSKTIKKKSSSSPSPRSNKKSNOHDGKEKLEKNST     SSTIKKKESSSPSPSRSNKKSNOHDGKEKLEKNST     SSTIKKKESSSPSPSRSNKSNOHDGKEKLEKNST     SSTIKKKESSSPSPSRSNKSNOHDGKEKLEKNST     SSTIKKKESSSPSPSSRSNEKKEGER     HISERRERERSTSNKKSSSNOHDGKEKLEKNST     SSTIKKER     STIKKTIKKSSSSSPSPSRSNEKNEKERER     HISERRERERSTSNKKSSSNOHDGKEKLEKNST     SSTIKKKESSSPSPSRSNEKNEKERER     HISERRERERSTSNKKSSSNOHDGKEKLEKNST     SSTIKKTIKKSSSSSPSPSRSNEKNEKERER     HISERRERERSTSNKKSSSNOHDGKEKLEKNST     SSTIKKTIKKSSSSSPSPSRSNEKNEKERER     HISERRERERSTSNKKSSSNOHDGKEKLEKNST     STIKKTIKKSSSSSSPSPSPSRSNEKNEKNEKER     STIKKTIKKSSSSSSSSSBSSPSPSPSRSNEKNEKNEK     STIKKTIKKSSSSSSSSSBSSPSPSPSRSNEKNEKNEK     STIKKTIKKSSSSSSSSBSSRSRRRRRSSSSSS     SSTIKKTIKKSSSSSSSSSBSRRRRRSSSSSSSSSSSSSSSSSSSSS			<u> </u>			·	
ATLGEPGPPLMGNVDPSKIDEIRRITVYYGH. NSQTTTADQLLEFEKQVGEWKFYRMAGDET QPTRRAFVEFAQQONSVPRALAFNGVMFGDRP LKNNISNATVEPEMTPQAAKELEVMKR VREAQSFISAAIEFGWLHSTSL.CNDFLGGFRR RMMTRPACTICGTFHLCLINWOL-LPAYTA KVFPPFRVWKEGYEKRRUSRSHTRSKSRSSKS SHSKKKSSSKHRSKSRSHRSKSRSSKSSK SHSKKKRSSSKHRSKSRSSKHRSKSRSKSKSK SHSKKKSSKHRKSRSTSHRSKDRKKDTREKI KEKERVKEKDREKEREREKEREKEKERGK KORDKGERKDREKOKKEKDERERREKHEKD ROKEKERGDKEKEREREKEREKER KEKERVKEKDREKERERSKERSSSSK KRITSKTIKKSSRSPSPRSNKKDKKREERER HISRERBERSTSKRSKSSSKERRSSSSSS PRITSKTIKKSSRSPSPRSNKKDKKREERER HISRERBERSTSKRKSSSSTORMSKLSKNTLTUTVLNSQDP WFFFYPSPLQIPSSFRNGLNDWYHHSQKC PULDGAYVKK  376  1725 A 3192 415 101 AHSSHOTRALIGEFGWDIRHPPLSPNLALSG FPPNLKSLRGTHFSSVKKNTLTUTVLNSQDP WFFFYPSPLQIPSSFRNGLNDWYHHSQKC PULDGAYVKK KRTPPPPAFFVLEFGFLAGGAGLAGDPPAS ASQSAGITGVSHTWFKNHLIFYACLVTRSKRI K KTCYTSRGSPLSPQSSIDSELSTSELEDDSISM KKUNDIGNARALGEESKROPYASTSAS VSRHSSSVSLSSGKGTCSDQEYDQVSLEDGE EFDHLPPGPRLPRCSFFGRGIPSGTFSSIREC RSFSSQYFSNNYQQQYYSPQAGTPDQV NRTNGDKPPKKYAPSGRQYASTSAS VSRHSSSVSLSSGKGTCSDQEYDQVSLEDGE EFDHLPPGPRLPRCSFFGRGIPSGTFSSIREC RSFSSGYFSNNYQQQYYSPQAGTPDQV NRTNGDKPPKKYAPSGRQYKATSAS VSRHSSWSLSSGKGTCSDQEYDQVSLEDGE EFDHLPPGPRLPRCSFFGRGIPSGTFSSIREC RSFSSGYFSNNYQQQYYSPQAGTPDQV NRTNGDKPPKKYAPSGRAFXAASGBAGRS ALPRESLANGSNLPRSKLISRFVGWG- PFKVNRYTFINRQPLRVCHRALELTGRWLIN LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPDHFTHIRIDEAGHCMEFESLVAIA LVSAQPPTHFTLDFDNQLYVFGELQACA DVVDREEFCRWAGLEPQGPTIFHGVMCKL ERGONSPSFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPTRAQUERFTCHOMICKL ERGONSPSFNPEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPTRAQUERFTCHOMICKL ERGONSPSFNPECAATVTSYLKLLAPSSKK GKARLSPRSVGVISPTRAQUERFTCHOMICKL ERGONSPSFNPECAATVTSYLKLLAPSSKK GKARLSPRSVGVISPTRAQUERFTCHOMICKL ERGONSPSFNPECAATVTSYLKLLAPSSKK GKARLSPRSVGVISPTRAQUERFTCHOMICKLOP EKSSSHISSRREPTPLAATNRAAALFEPLTFOD SNLRVWINGORICPACTTTSCHS  379 1729 A 3205 B 3205 B 3205 B	374	1724	Α	3187	191	1815	
NSQTTTADQLLEFFKQVGEVKFVRMAGDER    QFTRRAFVEFADQNSVPRALARNGVMFGDRP    LKINHISNNAIVXPPEMTPQAAAKELEEVKMK    VREAGSFISAALPEGWLHSTSLCADFLGGFVRR    RMYRE*APCTICGTFHLCLINWDL*LIP*ATA   K*FFPPRVWKEQ*KKRRRSRSHRRSKSSSSK   SHKKRKSQSKHRSRSHRRSKSRSKSSSSK    SHKKRKSQSKHRSRSHRRKSKSRSSKKDRRSKK    KKEKEVBKDREKEREKKEREKKRGKN    KKREVVEKDREKEREKKEREKEKERGKN    KORDKERKDREKDKEKOREKEREKKEREKKORK    KREKEVBKDREKEREKKEREKKERKOK    KREKEVBKDREKEREKKEREKKERKOK    KREKEVBKDREKEREKKEREKKERKOK    KREKEVBKDREKEREKKEREKKEREKKOK    KRETPRSYNASRRSSSSRERRRRSRSSSS    PRISKITIKKKSSSSPSPRSNRKKKKEERD   HISBERERERSTSMRKSSNDROGKEKLEKNST    SHITTEN     SHITTEN		'	1		]		AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI
QPTRFAFVERADQNSVPRALAFNOYMFGDRP LKNNISNNALVKPPEMTPQAAAKELEEVMKK VREAQSFISAAIEPGWILISTSLCNDFLGG*RR RMYME*APCTICGTFHLCLINWID-LIP*ATTA K*PFPPRWKEO*PKKRRISSHTRSKSSSSK SHERKENGOSKHRSSHTRSKSSSSSK SHERKENGOSKHRSSHTRSKSSSSSK SHERKENGOSKHRSSHTRSKSSSSK SHERKENGOSKHRSSHTRSKSSSSK SHERKENGOSKHRSSHTRSKSSSSK SHERKENGOKHRSKHEIDSKRKDENGKRKDK KSKTPRSYNASRKSSSSERREREKEREKEREKER KEKERVERCDREKGEREKEREKEREKEREK KEKERVERCDREKGEREKEREKEREK KSKTPRSYNASRKSSSSSFRRRRRSRSSSSSS BRTSKTIKRKSSSSSPSPRSENKEDIGKREKDK KSKTPPRSYNASRKSSSSSFRRRRRRSSSSSSS PRTSKTIKRKSSSSSPSPRSENKEDIGKREKDK KSKTPPRSYNASRKSSSSSSERRRRRSRSSSSSS PRTSKTIKRKSSSSSSPSSRNKDIKKERERD HISERREREGSTSMKCSSNADDGKEKLEKNST S  375 1725 A 3192 415 101 AHSSHOTRAILQEFQWDIRHPPLSPNLALSG PFPALKKSLRGTHFSSVKKTTLTWLNSQOP WEFFYP*SEDQLOPSSFRNGLICNDWYHHSQKC PDLDGAVVKK DUDGAVVKK DUDGAVVK  GV*WCDLGSPQPPPFGKQFCLGRSSSWDYR HVPPHPANFVFLLETGFLHAQQAGLIGDPPAS ASQSAGTTOVSHTWFKHILETY ACUVTRSKRI K KTGYTSRGSPLSPQSSIDSELSTSSLEDDESTSS LTRAIN STREET STSELEDDSTSS VSRHSSSVSLSSGKKGTCSDQFYDQYSSLEDEE EFDHLPPPQPAFPLRCSPPGRGPHSGTSTSELEDDSTSS VTVRNSQSFDSSLHAAGRGISRIQGCTSPQGU QHRYBSVGHPPSIRQPLKATAYVSFTVQGSS NMPLSNGLQLYSNTGIPTPMKAAASGIMGRS ALPEPSLANGOSLLPSKLANGTH*9GHSSP VTVRNSQSFDSSLHAAGRGISRIQGCTSPQGU QHRYBSVGHPVSIRQPLKATAYVSFTVQGSS NMPLSNGLQLYSNTGIPTPMKAAASGIMGRS ALPEPSLANGOSLPSKLANGATS*PGTUCPKPKP  TOP OF THE CONTROL	ļ	ì	1	l	Ĭ	}	ATLGEIPOPPLMGNVDPSKIDEIRRIVYVGNL
LKNNISNNAIVKPPEMTPQAAAKKLEEVMKR VREAQSFISAAEPOWLISTSICNDFLGCFPRR RMYRE*APCTICGTPHICLINWDL*LP*AYTA K*FPPRVWKSQ*KKRRSISICNDFLGCFPRR RMYRE*APCTICGTPHICLINWDL*LP*AYTA K*FPPRVWKSQ*KKRRSRSSRSRSQKDRRSSK SHSRKRSQSKHRSRSHNSRSRQKDRRSK SHSRKRSQSKHRSRSHNSRSRQKDRRSK KERRYKEKDRKEKEREKEKEREKEKERK KERKEKEKEREKEKERKEKER RDEKEKERGDKEKEREKEKEREKEREKER RDEKEKERGDKEKEREKEKEKEREKEREKER RDEKEKERGDKEKEREKEKEKEREKEREKEREKER RDEKEKERGDKEKEREKEKEKEREKEREKER RSRSSSRSSRSTRTRTRSTRSSSSSS PRTSTITIKTRSSSPSPRSRKKOKTEKERD HISERRERESTSMRKSSNPSDGKEKLEKNST S  375 1725 A 3192 415 101 AHSSHOTRAIL QEFQWDIRPPLSPNLALSG PFPNLKKSLRGTHFSSVKKTTLTWLNSQDP WF#FFYP*SPDLQPSSFRNGLNDWYHHSQKC PDLDGAVVK PDLDGAVVK PDLDGAVVK VPPHPANFVLLETGFLHAGQAGLAGDPPAS ASQSAGTGVSHTWFKNHLITAGLAGVAGLAGDPPAS ASQSAGTGVSHTWFKNHLITAGLACVTRSKRT K  377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQVASTSAS VSRTISSSVALSSGKKGTGSDQVOYSLEDEE EFDHLPPPQPRLPRCSPFQRGPHSOTFSSREC RRSFSSQYFPSNNYQQQYYSPQAGTPDQP NRTNGDKFPKKYA*PSPDAGVYSLCDPEE EFDHLPPPQPRLPRCSPFQRGPHSOTFSSREC RRSFSSQYFPSNNYQQQYYSPQAGTPDQP NRTNGDKFPKYA*PSTPGNSLQPKYPKY- SSLSTLRDGNWRDGCY VYKRYSVLGPHEXATAYVSPTYQGSS NMPLSNGLQLYSNTGIPTPNKAAASGINGRS ALPPSLANDSNILPSSKLQNSTLOPPFTQL VKGFYHKVALDRVKLSFSMSLLSRFVQWG* PPKVNFYTFRAPQELXYQHEXLTGYMLTW PMLFFVAPROVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPTHEPPGTGKTVT LVEARCQVYKHLPSSTYRLLAPSSKIMMVEBIPPCON WDAKKGEYVPFAKKLLQFRYLTTLTTAGR LVSAQPFIDHFTHIFIDEAGICMEESISPCOL QRLKYHLPSSTYRLLAPSSKIMGNEDEIPCON WDAKKGEYVPFAKKLLGFRYLLTU LLPEARCQVYKHLPKSTYLLLLAPSSKIMGKD LLPEARCQVYKHLPKSTYLLLLAPSSKIMGK GRARLSPRSVQVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKYVTCOSTVTPCLFCAPCIPLP ETSSSFHISSPRPRTFAALNRARAFLEPLTTPOL SNLRWWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVULWPPFFL*GSFFPTGHTXCVXIK	]	ļ	}	1		ł	
VREAGSFISAAIEPGWILHSTSLCNDFLGG*PRR	1	i		1		ļ	
RMYRE*APCTICGTFILCLIRNVDL*L*ATA*			1	1			LKINHSNNAIVKPPEMTPQAAAKELEEVMKK
K*PFPPRWKEQ*KKRRRSSHTRSKSSKSSK SHSRKRRSQKBRRSSK  SHSRKRRSQKBRRSSHTRSKDKRDTREM  KEKERVKEKDREKEREKEREKEKERKEKERKKDK KEKERVKEKDREKEREKEREKEKEREKEREKEREKEREKEREKERE				i	ł	1	
SHSRKRSSSKHRSRSHSRDKRKDTREKI   SHSKRERKSSSSSHSRDKRKDTREKI   KEKERVKEKDREKEREKEREKERCKER   KEKERVKEKDREKEREKEREKERCKER   KDRDKEREKDOKEKDREKEREKEREKERCK   KDRDKEREKDOKEKDREKEREKEKERCK   KDRDKEREKDOKEKDREKEREKEKED   KRSTPFRSYNASRSRSSSSERRKRSKSSSSS   PRTSKTIKKSSSSSPSTRKKSSRSSSSSS   PRTSKTIKKSSSSSPSTRKKSSSSSSPSTRKSSSSSSERT   KRSSSSSSPSTRKSSSSSSPSTRKSSSSSSPSTRKSSSSSSSSPSTRKSSSSSSSSSS	1	1	1	1	ì		
SPHIKKRSKSRERKKSRSSHSINDKRDTREKG KEKERVEKDREKERREKEKEKERGKN KORDKERELDEKKEKEKEKEKEKEKEKEK KEKERVEKENDREKEREKEKEKEKEKEKEK RORDKEKEKOPKEKEREKEKEKEKEKEKEK RORDKEKEKOPKEKEREKEKEKEKEKEK RORDKEKEKOPKEKEREKEKEKEKEKEK KORDKEREKOPKEKEREKEKEKEKEKEK KERTPFRSYNASKRSSSSBERRRRSSSRSS PRTSKTIKKSSRSPSPSRSNKKDKKREKER HISEREERERSTSMRKSSDROGKEKLEKNST S  375 1725 A 3192 415 101 AHSSHQTRAILQEFQWDIJRHPPLSPNLALSG PPLIDGAYVKK SPPLIKKSLROTHFSSVKKITLTWLNSQDP WF/FFYP*SPDLQJPSSFRNGLNDWYHHSQKC PPLDGAYVKK PPLDGAYVKK  376 1726 A 3199 931 418 GV*WCDLGSPQPPFGFKGFCLGRSSSWDYR HYPPHPAHPAFVFLETCFFLHAGQAGLAGDPAS ASQSAGITGVSHTWFKNHLIFYACLVIRSKRI K  377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQJMARLQEESLRQDYASTSAS SYRHSSSVSLSSKKGTCSDQFYDQYSLEDBE EFDH.PPPQPRLPRCSPFQRGJPHSQTFSREC RSSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDK/PFKKYA*PSPDAKYNGH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCUPSPGQL QRVHISVGHPVSIRQPLXATAYVSPTVQGSS NMPLSIGLQLYSNTGJPTPNKAAASGIMGRS ALPPRSLANGSNLPRSKJQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VFGVTESRFSVLRDDHLFALLSSETHQEDPIT VFGGYHKVELDRVKLSSFMSLLSRFVQWG* PFKVNFY/TFNRQPLRVOQHRALELIGRWLLW PMLFPVAPRDVPLLPSDKJKLTDTNSLESN EQLQAMRHIVTGTTRPAPYHFGPPGTGKTVT LVEAKQFVKHLVKLATDRSLESN PMLFPVSVRAPRDVPLLPSSTMSLLSRFVQWG* PFKVNFY/TFNRQPLRVOQHRALELIGRWLLW PMLFPVSNFYLDPHTHTIFIDEAGHCMEPSSLVAIAG LMSVKETGDPGGQLVLAGDPRQFDIPOQPTDPQ PTTKLLRNYRSHFTILDIPNQLYYEGGLQACA DVVDRERFCRWAGILPRQGFPIHFQMGKD ERGSNSPSFFNPEAATVSTYLKLLAPSSKK GKARLSPRSVGVISPYRKQUPEKNCTIKLDR ERGSNSPSFFNPEAATVSTYLKLLAPSSKK GKARLSPRSVGVISPYRKQUPEKNCTIKLDR ERGSNSPSFFNPEAATVSTYLKLLAPSSKK GKARLSPRSVGVISPYRKQUPEKNCTIKLDR ERGSNSPSFFNPEAATVSTYLKLLAPSSKK GKARLSPRSVGVISPYRKQUPEKNCTIKLDR ERGSNSPSFFNPEAATVSTYLKLLAPSSKK GKARLSPRSVGVISPYRKQUPECTCHLPCAPTCPLP ETSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRRACLTNTSCHS SNLRVWDGIRRACLTNTSCHS SNLRVWDGIRRACLTNTSCHS SNLRVWDGIRRACLTNTSCHS SNLRVWDGIRRACLTNTSCHS SNLRVWDGIRRACLTNTSCHS SNLRVWDGIRRACLTNTSCHS			1		1	į	K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSSK
			1		ł	1	
SURDKEREKDREKDEKENDERFREKEHEKD   RNKEKEKGONKEKEREKDEK   RNKEKEKGONKEKERSKEIDEKEKKOK   RNKEKEKGONKEKERSKEIDEKEKKOK   RNKEKEKEONKEKERSCHORKEKEKEN   RTSKTIKKKSSRSPSTRSRNKKKKKKEKERD   HISERRERESTSTMRKSSNDROKEKELEKNST   S   S   S   S   S   S   S   S   S	ł		i	1	1	1	
RINGEKEKEQDKEKEREDEKREKDEN   RISSEDEKREKDEK   RISSEDEKREKEKDEK   RISSEDEKREKDEK   RISSEDEKREKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKREKTEKDEK   RISSEDEKRETEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSEDEKRETEKTEKDEK   RISSED	1	1		1		ì	
SSRTPPRSYNASRRSSSSRERRRSRSSSSRERRRSRSSSSRER   RTSKILKRLSSRSPSSPRRKKLKKRERD   HISERRERESTSMRKSSNDRDGKEKLEKNST   S	ł	Į.	ı	l l			
PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD   HISERRERESTSMRKSSNDRDGKEKLEKNST     S	1	1	1				
HISERRERSTSMRKSSNDRDGKEKLEKNST   S			1	Ì	}	1	
S	]			1			PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD
1725   A   3192   415   101	1	1.	1				HISERRERERSTSMRKSSNDRDGKEKLEKNST
PFPPLKKSLRGTHFSSVKKITLTWLNSQDP   WF/FFYP*SPDLQIPSSFRNGLNDWYHHSQKC   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PDLDGAYVKK   PVPPPANIVFILLETGFLHAGQAGLIGDPPAS   ASQSAGITGVSHTWFKNHLIFYACLVIRSKRI   K   KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM   GYKLQDLTDVQIMARLQEESLRQDYASTSAS   VSRHSSSVSLSSGKKGTCSDQEYDQVSLEDEE   EFDHLPPPQPRLPCSSFQRQIPHSQTFSSIREC   RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP   NRTNGDK/PKKYA*PSPDAKYNCH**QHISSP   VTVRNQSFDSLHGAGNGISRIQSCIPSPGQL   QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS   NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS   ALPRPSLAINGSNLPRSKLAQPVRSFLQPPKPL   SSLSTLRDGNWRDGCY   SSLSTLRDGNWRDGCY   SSLSTLRDGNWRDGCY   VFQVTESRPSVLRGDHLFALLSETHQEDPIT   YKGFVHKVELDRVKLSFSMSLLSRFVGWG*   PFKVNFY/TFNRQPLRVQHFALELTGRWLLW   PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP   EQLQAMRHIVTGTTRPAPIITGPPGTGKTVT   LVEAKQVVKHLPKAHLACAPSNSGADLLC   QRLRYHLPSSIYRLLAPSRDIRMYPEDIKPCCN   WDAKKGEYVFPAKKKLQEYRVLITTLITAGR   LVSAQPPIDHFTHHIPIDAGHCMERESLVAIAG   LMEVKETGPPGGQUVLAGDPRQLGPVLRSPL   TQKHGLGYSLLERLLTYNSLYKKQPGGYDPQ   FITKLLRNYRSHPTILDIPNQLYYEGELQACA   DVVDRERFCRWAGULPRQGFPIIFHGVMGKD   EREGNDSPSFFNPEEAATVTSYLKLLAPSSKK   GKARLSPRSVQVISPYRKQVEKIRYCITKLDR   ELRGLDDIKKDLKVTCCSTVTPCLPCAPTCPLP   ETSSSFHSSRPRPTPAALNRARALPELTFGD   SNLRVWDGIRKPACLTNTISCHS   SNLRVWDGIRKPACL	1	1.		1.			
FFPPILKSLGTHFSSYKKITLTWLNSQDP   WF/FFYP*SPIQIPSSFRNGLNDWYHHSQKC   PDLDGAYVKK	375	1725	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG
PDLDGAYVKK   GV*WCDLGSPQFPPFGFKQFCLGRSSSWDYR   HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS   ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI   KVPPHPANFVFLLETGFLHAGQAGL\GDPPAS   ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI   KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM   GYKLQDLTDVQIMARI.QEESLRQDYASTSAS   VSRHSSSVSLSSGKKGTGSDQSYDQVSLEDEE   EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC   RSPSSQYFPSNNYQQQYYSPQAQTPDQQP   NRTNGDK/PPKKYA*PSPPAKYNCH*QHSSP   VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL   QHRVHSVGHFPVSIRQPLKATAYVSFTVQGSS   NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS   ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL   SSLSTLRDGNWRDGCY   SSLSTLRDGNWRDGCY   VSQTTESRFSVLRGDHLFALLSSETHQEDPIT   YKGFVHKVELDRVKLSFSMSLLSRFVGWG*PFKVNFY/TFNRQPLRVQHRALELTGRWLLW   PMLFPVAPRQVPLLFSDVKLKLYDRSLESNP   EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT   LVEAIKQVVKHLPKAHILACAPSNSGADLLC   QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN   WDAKKGEYVPPAKKKLQBFVLJTTLTTAGR   LVSAQFPIDHFTHFIIFIDEAGHCMEPESLVAIAG   LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL   TQKHGLGYSLLERILTYNSLYKKGPDGYDPQ   FITKLLRNYRSHFTILDIPNQLYYEGELQACA   DVVDRERFCRWAGUPRQGFPIIFHGVMGKD   EREGNSSFFNPEAAATVTSVLKLLAPSSKK   GKALSPRSVGVGATVTSVLKALLAPSSKK   GKALSPRSVGVGATVTSVLKALLAPSSKK   GKALSPRSVGVGATVTSVLKATALPEPLTFGD   SNLVWDGIRKPACLTNTSCHS   SNLVWDGIRKPACLTNTSCHS   SNLVWDGIRKPACLTNTSCHS   STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STAT	1 3.3	1	1	1		1	F\FPNLKKSLRGTHFSSVKK\TTLTWLNSQDP
PDLDGAYVKK   GV*WCDLGSPQFPPFGFKQFCLGRSSSWDYR   HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS   ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI   KVPPHPANFVFLLETGFLHAGQAGL\GDPPAS   ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI   KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM   GYKLQDLTDVQIMARI.QEESLRQDYASTSAS   VSRHSSSVSLSSGKKGTGSDQSYDQVSLEDEE   EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC   RSPSSQYFPSNNYQQQYYSPQAQTPDQQP   NRTNGDK/PPKKYA*PSPPAKYNCH*QHSSP   VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL   QHRVHSVGHFPVSIRQPLKATAYVSFTVQGSS   NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS   ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL   SSLSTLRDGNWRDGCY   SSLSTLRDGNWRDGCY   VSQTTESRFSVLRGDHLFALLSSETHQEDPIT   YKGFVHKVELDRVKLSFSMSLLSRFVGWG*PFKVNFY/TFNRQPLRVQHRALELTGRWLLW   PMLFPVAPRQVPLLFSDVKLKLYDRSLESNP   EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT   LVEAIKQVVKHLPKAHILACAPSNSGADLLC   QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN   WDAKKGEYVPPAKKKLQBFVLJTTLTTAGR   LVSAQFPIDHFTHFIIFIDEAGHCMEPESLVAIAG   LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL   TQKHGLGYSLLERILTYNSLYKKGPDGYDPQ   FITKLLRNYRSHFTILDIPNQLYYEGELQACA   DVVDRERFCRWAGUPRQGFPIIFHGVMGKD   EREGNSSFFNPEAAATVTSVLKLLAPSSKK   GKALSPRSVGVGATVTSVLKALLAPSSKK   GKALSPRSVGVGATVTSVLKALLAPSSKK   GKALSPRSVGVGATVTSVLKATALPEPLTFGD   SNLVWDGIRKPACLTNTSCHS   SNLVWDGIRKPACLTNTSCHS   SNLVWDGIRKPACLTNTSCHS   STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STATUTT STAT	ļ	}			1		WF/FFYP+SPDLQIPSSFRNGLNDWYHHSQKC
HYPPHPANFVFLLETGFLHAGQAGL/GDPPAS ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI K  377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVLSSGKKGTCSDQEYPQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSSVLSSGKKGTCSDQEYPQYSLEDEE FDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDKPPKKYAPSPDAKYNCH**QMSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS AALPRSSLANGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPFGTGKTVT LVEAIKQVVKHLPKSHTLAPSSRIMVPEDIKPCCN WDAKKGEYVFPAKKKLQBYRVLITTLTTAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITIKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRSTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS SILRVWDGIRKPACLTNTSCHS STPP 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1		1	1	Ì		PDLDGAYVKK
HYPPHPANFVFLLETGFLHAGQAGL/GDPPAS ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI K  377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVLSSGKKGTCSDQEYPQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSSVLSSGKKGTCSDQEYPQYSLEDEE FDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDKPPKKYAPSPDAKYNCH**QMSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS AALPRSSLANGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPFGTGKTVT LVEAIKQVVKHLPKSHTLAPSSRIMVPEDIKPCCN WDAKKGEYVFPAKKKLQBYRVLITTLTTAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITIKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRSTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS SILRVWDGIRKPACLTNTSCHS STPP 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	376	1726	1 A	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR
STOP	370	1,20	1.	7.77		1	HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS
STOP	1	1	1	1	1	}	ASOSAGITGVSHTWPKNHLIFYACLVIRSKRI
GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSGKKGTCSDQEYDQYSLEDEE EFDHLPPPOPRLPRCSPFORGIPHSQTFSSIREC RSFSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRESLAINGSNLPRSKLAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHK VIELDRVKLSFSMSLLSRFVGWG* PFKVNFYTFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIEPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGQLVLAGDPRQLGPVLRSPJ TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGYMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRTPAALNRARALFEPLTPGD SNLRVWDGIRKPACLTNTSCHS  1729 A 3206 432 130 PKAAPSVXLWFPFIL*GSFKPTKGHTXCVXIK	1	1				1	
GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSGKKGTCSDQEYDQYSLEDEE EFDHLPPPOPRLPRCSPFORGIPHSQTFSSIREC RSFSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRESLAINGSNLPRSKLAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHK VIELDRVKLSFSMSLLSRFVGWG* PFKVNFYTFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIEPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGQLVLAGDPRQLGPVLRSPJ TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGYMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRTPAALNRARALFEPLTPGD SNLRVWDGIRKPACLTNTSCHS  1729 A 3206 432 130 PKAAPSVXLWFPFIL*GSFKPTKGHTXCVXIK	277	1727	+ <u> </u>	3201	274	1285	KTGYTSRGSPLSPOSSIDSELSTSELEDDSISM
VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE EFDHLPPPQPRLPRCSFFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QHISSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VPGVTESRFSVLRGDHLFALLSSETHQEDPIT YKGFVHK VIELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEVRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LWSQVFPIDHFTHIFIDEAGHCMEPESLVAIAG LWSQFVFPAKKKLQEVRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LWSQFPIDHFTHIFIDEAGHCMEPESLVAIAG LWSQFPIDHFTHIFIDPNQLYYEGELQACA DVVDRERFCRWAGLLPRQGFPIIFHGVMGKD ERGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRPTPAALNRARRALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	3//	1/2/	A	3201	12/4	1203	
EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH***QH\SSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRSLANGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFYTFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRRAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQBYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYFGELQACA DVVDRERFCRWAGULPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK				l.			
RRSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDKPKKYA*PSPDAKYNCH**QHNSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLANGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFYJTFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIFNQLYYEGELQACA DVVDREFFCRWAGULPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1 .	ł	ļ	1	1		
NRTNGDK/PPKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVIHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGULPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTTKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRYWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK				}	}		RRSPSSOYFPSNNYOOOOYYSPOAQTPDQQP
VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHK-VLDRVKLJSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRVCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRRPACLTNTSCHS  179 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1	1					NRTNGDK/PPKKYA*PSPDAKYNCH**OH\SSP
QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLANGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRFTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK		ì		1	1		VTVRNSOSFDSSLHGAGNGISRIOSCIPSPGOL
NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1				İ		OHRVHSVGHFPVSIROPLKATAYVSPTVOGSS
ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QFRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGJLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPSPTPAALNRARALFEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1	1		1	ľ		NMPL SNGLOLYSNTGIPTPNKAAASGIMGRS
SSLSTLRDGNWRDGCY  378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG'LLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK				1		, '	AT PROST AINGSNI PRSKIAOPVRSFLOPPKPL
378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG'LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK				İ	1		
YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG'LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK		<del> </del>	<del>                                     </del>	2000	112	1700	
PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRFAPYIIFGPFGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	378	1728	A	3202	112	1/07	VKCENHKIVEL DBAKL GESWELT SEEAGMC+
PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG'LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK		1			1		
EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDP FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG'LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1		ļ		
LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPV FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG'LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			- [				ENT O YMPHINACALDD Y DAILE GDDULUK ANA
QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDREFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1	1	1		EQUANKALY IOTIKAA III. OCTOTOKI VI
WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1		1	ĺ	Į.		ORI DARE DOSINELLA DODDINA CORDIVACALI
LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK		1	1				
LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	İ			Į.	1	1	
TQKHGLGYSLLERLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1				ł	1	
FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1		İ				LWEAKEI GDLAGOTATE AND COLORS TO THE COLORS
DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1		1		TQKHGLGYSLLEKLLTYNSLYKKGPDGYDPQ
EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1		1		FITKLLRNYRSHPTILDIPNQLYYEGELQACA
GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			- 1	1	1	1	DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD
ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1	i	1		EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK
ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS  379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1			ł	GKARLSPRSVGVISPYRKQVEKIRYCITKLDR
SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			}		1		
379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	İ	1	1	İ	1	1	ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD
379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1	}		i	1	ļ	SNLRVWDGIRKPACLTNTSCHS
	370	1729	A	3206	432	130	
	13,7	1	1			1	*LSTREAXDSXPGRQIAXXRQGGKVETTTAL

						Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, F=Fronine,
uence	1	Į.	914 '	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ	}		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1	l	peptide	1	/=possible nucleotide deletion, \=possible
	1	ļ		sequence		nucleotide insertion
	<del> </del>	<del> </del>	<del>                                     </del>			XKQSNNKGTRASSYXEPDAXEQWKFPHKKL
	1	l	1			QLPGXTHE
380	1730	A	3207	187	507	GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS
300	1730	1 ~	1 3207	1 ***	1	PEORDHPAPHKOIEAGOGLPGPQAWGG*KGP
	1	j	}	ļ	1	AXLLPGPGGGPGPVASLEARAQASSGVTPNG
	1	1	1			GGRTYPYPTFSSGE
201	1531	<del>  .      </del>	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/
381	1731	A	3223	1 *	040	EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI
	1	Ĭ	1			FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS
	Į	Į	1	1		KAWRASQMMTFFIFLLFFPSFTGVLCTLAITI
		1	l	<b>\</b>		WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP
	]	1	1	ł		WKLKPSADCGPFKOLPLFIRSTTSWIDTLSTKI
		1	ļ			GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY
	1		]	ì	1	WQITEGRKIMIRLLHEQIINEGKDKMFLIEKLI
	1	1	l	1	1	KLQDMEKKANPSSLVLERREVEQQGFLHLGE
			ļ	]	_	HDGSLDLRSRRSVQEGNPRA
382	1732	A	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHHRHHQGHNS
302	1			i		LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF
ł	1	1	1			LFSNACL
383	1733	TA.	3241	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD
202	1733	1	3211	10.2	1	KVTMLWNKKATAVLVIASTDVDKTGASYYG
		1		1		EQTLHYIATNGESAVVQLPKNGPIYDVVWNS
1	1	1		ł		SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG
	Į	1	1	l		PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK
Ì		1	ì	1		VWNVKNYKLISKPVASDSTYFAWCPDGEHIL
ļ	1		<b>\</b>	1		TATCAPRLRVNNGYKIWHYTGSILHKYDVPS
1 .	1	1	l l	1	ł	NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP
Į.	1	1	ì			NAELWOYSWOPPLINGFFARIII I QAYI SEYI
	İ	1	1			NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ
	1	ì	ì		1	NMKPQSGNDKPLSKTALKNQRKHEAKKAAK
}		}	1	j		QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP
		1				EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK
			1		ţ	NQLEKIQKETALLQELEDLELGI
384	1734	A	3242	3	678	IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP
1 50.		1		ł		RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP
		1		1		LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS
	1	1	1		ļ	ASLVRATVRAVSKRKLOPTRAALTLTPSAVN
1	1	1	- 1	ı	i	KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL
		i i		Ĭ	†	EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL
İ	1	1	1			GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES
}	1	1	}			FNI
			2012	12100	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL
385	1735	A	3243	3190	1 00-4	KEEEILPEPGSETPTVASEALAELLHGALLRR
						GPEMGYLPGPPLGPEGGEEETTTTITTTTTTT
1						TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL
1	}	1		l	- {	GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL
		1		1		ALL ACCORDOL ADDIT AND A CECOM DOLL
		-		}	1	VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT
	1	i	1	(	1	NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP
1		i		- [		PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG
			1	1	į	EETLICLNGTRPSWNGETPSCMASCGGTIHNA
	1	1	1	1		TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL
		{	1	l	. [	HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS
	ļ.	İ			1	DMDDVPERGLISDAQSLYVELLSETPANPLLL
1		- 1				SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG
	1	-	1	Ì		ALATFSCLPGYALEPPGPPNAIECVDPTEPHW
	1	J	1	ŀ	1	NDTEPACKAMCGGELSEPAGVVLSPDWPQS
	i		ı			
				1	<u> </u>	
						YSPGODCVWGVHVQEEKRILLQVEILNVREG
						YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAOLRGPOPRRRLLSS
	į					YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR
			-			YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAOLRGPOPRRRLLSS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nuci-	peptide		in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	!			peptide	Coducinos	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			<b></b>	sequence		MTCADPGEIANGHRTASDAGFPVGSHVQYRC
	· '	1	1			LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC
	i	l	ļ	l .		ALKYEPCLNPGVPENGYQTLYKHHYQAGESL
	1	}	j			RFFCYEGFELIGEVTTTCVPGHPSQWTSQPPLC
		<b>l</b> .	1			KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG
	1		1		1	SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF
		)		İ	}	SNPLYEAGDTREYEVSI
386	1736	A	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT
500	1				1	RKLLMTWALEVAVVMKKSETYAPLFCLPSF
		}	}		1	HKFCKGLLADTLVEDVNICLQACSSLHALSSS
		l		1		LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL
	1	l	Ì	1	(	LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP
						SNTFHPQDFSD/VISFILYGNSHRTGKDNWLE
		1				RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW
	İ			1		AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR
	1	1		1		SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP
	1				1	KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA
			_	1	1	GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI
		· ·	1	· ·		MMVVEALCELHCPEAIQGIAVWSSSIVGKHL
		1		ł	]	LWINSVAQQAEGRFEKASVEYQEHLCAMTG
					1	VDCCISSFDKSVLTLASAGCKSASLKHCLNGE
				1		SRKSVLSKPTDSSPEVINYLGNKACECYISTA
				1		DWAAVQEWQNAIHDLKKSTSSTSLNLKADF
	Ì		1.	1		NYIKSLSSFESGKFVECTEQLELLPGENINLLA
		·		1		GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYOVOTEI*NSIQHIMA\SKKLSRF
301	1737	12	3233	1 500		LKYVHNL*AENYKTLMK*INEDLNKQRDVPY
		1	1			S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET
		1		}	1	NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG
1	1			1		FELLDSSDLPASASKSAGITCMSHHARTLSLK
			1	1		*WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI
						LTRLETOMINADYQNKLTLDYLLTTDREVYE
	1	ł	1	1		PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV
i						PVQV*HGFDPEAMFR
390 -	1740	A	3270	2	372	GRCHDQNKGKS\DGPDAQAEACGGESTYQEL
1		1				LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV
1	1	- 1				YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
					<del></del>	FFOEMLDIMKAISDMMGKCTYPVLKEDAPRQ
391	1741	A	3273	1	187	HVETFFQ\'EELTRSQEGMKLGENFLMFAMPP
		j		1		DDSKESKGK*FFQEMLDIMKAISDMMGKCTY
		1				PVLKEDAPRQHVETFFQVGINQKSRGHEVRR
1		- 1		1	ļ	KFPDVCHAPR
					- 601	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK
392	1742	A	3281	901	521	RFSYLSLPSSWIDYRHVLPRQANFCIF/M*RRG
1					1	FTMLARMVSIS*PRDLPALASQSAGITGVSHH
}		1		1	}	APPOMDFTFALLCFALKGCLPRQKEGGTLNLI
		<del></del>	1000	705	3	RNRSVVPEFVLLGLSAGPQTQTLLFVLFVVIC
393	1743	A	3283	385	٦	LLTVMGNLLLLVVINADSCLHTPMYFFLGQL
	j	- 1	1	1	1	SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM
ļ		1	1	1		A*VFFVFATGGTESSLLAVMAYDRYVAIRTR
	ļ			1	1	
			1	l l	i.	1 (†
25.4	1544		2304	575	1054	G CTKCKADCDTCFNKNFCTKCKSGFYLHLGKO
394	1744	Ā	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC
394	1744	Ā	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKC
394	1744	Ā	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	peptide seq-		in USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
	uence	[	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine.
dence		]	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ	İ	}		peptide		/=possible nucleotide deletion, \=possible
1		}		sequence		nucleotide insertion
				1		RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
					}	KQQQ
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
		1			1	WICLSMVILTHSLKTFHRNWDWESEYTLFMS
1		1		1		ALKVNKNNAKLWNNVGHALENEKNFERAL
	<u> </u>					KYFLQATHVQPDDIGAHMNVGR
396	1746	Α	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
		ļ <u> </u>				IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	A	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL
	l	l	ļ			WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
i	ł	l		1	į	AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS
		ĺ			1	GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
398	1748	A	3300	1912	2768	GLT KORRWONIORKGPKRYIVIAGNSOSHOPMIFS
370	1/40	^	3300	1312	2100	MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
		1	1	1		STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
						LKGLYGNROLTARLELFTGRFKDWMVSMIV
	j	)		l		DREYSVAVEAVRLLILILKNMEGVLMDVDCE
	ļ	į.				SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
		i			1	RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
	}	1	İ		ļ	SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
1	ì	l		i		VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL
	L	İ	]			VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE
	1	1	1	1		SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
		1	[			CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
						PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L
]	1		]	1		STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ
	<b>!</b>	1	,			SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV
			ŀ			LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
	1				1	SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
	ĺ	ľ	[			N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
						TGAALAGSYPIWENENTLSWLPTFTYNFCLST
		1	1			PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
}	}		ł			NILPPNQTILISVEASISSSPIRNKWALHLITLLT
1	1					GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
	1		I			DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
1			1			TEKGGTCIYLQEECCFCVNESGIVHLAVRRLH
						DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
				1		LGPLIFLFLLMIGPCIFNLVSRFISQRLNCFIQ
		}		1		ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR P
400	1750	A	3303	12	453	THWRHSSGVPGSTTARRRRELEIATSDNOE
700	1.750	1	3303	1-	""	YYNRLCQEVTNRERNDQKMLADLDDLNRTK
		1	1		1	KYLEERLIELLRDKDALWQKSDALEFQQKLS
1						AEERWLGDTEANHCLDCKREFSWMVRRHHC
		]		1	1	RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
		1		i	1	CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP
		I		1		EKIVLRALKDSRAGMPEQDKDPGVQENPDD
				1		QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP
-				1		GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
		1		1	1	NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
		$\perp$		1		LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
1	1			1		PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF
	1			}		VVRTMCAVLGLVARQEDSGLRDHSVRVLISN
	_i					HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first armino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine O=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ FFFFIFLNIFLLAFSSPGSQPLLNSPPSFVCWSR GFMEMNGRGELVESLKRFCASTRLPPTPLLLF PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ VQRTLVSVTVSDASWVSELLWSLFVPFTVY QVRWLRPVHRQLGEANEEFALRVQQLVAKE LGQTGTRLTPA\DKAEHMKRQRHPR\LRPQS AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF CPHVAIGVFPERPWPKTGCCKTLTIHLILL*G GPVSFSCPE\DIHPRGT*VPTQQASGLPSFPSYG
403	1753	A	3307	44	447	PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL QERKQ\ALYEYARRFTERRAPGGLD DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS GGASAGLASSPECACGRSHFTCAVSALGECT
40.4	1561		2211	400	1	CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD CPPRECEED PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG
404	1754	A	3311	409	1	QDHVQNEEIYARVLDKFGSNFLSRDNADLGT AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS LLKGDLKGVKGDLKKPFDKAWKDYETKFAK IEKEKREREWR
405	1755	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG HPASKENDQMVDTIKNTTKVPIIWTYGDMVE PRPQMIRPAVGAKHKELWKILMALKKIK\IWE GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV MCVLLWALSLLQSILEWMFCSFLFSDVDSDN WCQILDFLTAVWLIFLI\LVLCGFTLVLLVRIIC GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F LLYWIEKDLDDL
407	1757	A	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC PHISFTGNICVYCPGGPDSDFEYSTGSYTGYED TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD KVEFIVMGGTFMALPEEYRDYFIRNLHDALS GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC MKRHLSDMLTYGCTRLEIGVQSVYEDVARD TNRGHTVKAVCESPHLAKDSGFKVVAHMMP DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP TLVIRGTGLYELWKSGRYKSYSPSDLVELVA RILALVPPWTRVYRVQRDIPMPLVSSGVEHG NLRELALARMKDLGIQCRDVRTREVGIQEIH HKVRPYQVELVRRDYVANGGWETFLSYEDP DQDILIGLLRLRKCSEETFRFELGGGVSIVREL HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL QGPYMVKMLK AIASPRAAGIRHELTSTMAAGKNKKLTKGGK
						KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG LKGRVFEESLADLQND\TDGYLLRVI*VAFTT ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK KGIRARILETLGMLLLLALLILGIVWVASALID NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK GTROPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGL\P*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY VKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342	74	2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRSVTM TDGLTADKVTRSDGCPTSTSLPRPDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEI
412	1762	A	3347	1	898	RRYEKFDKVLTALSHKLEPAVRSSEL  IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

			1		<b>TA</b> 12.3 1	A Alamina C-Custaina
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ł	in	nucleotide	location	
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ł	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Į.		peptide		/=possible nucleotide deletion, \=possible
		1	l	sequence		nucleotide insertion
			ļ	j		VMPLVRMPWKRAVVLLMLWFIGQAMWLAP
	1	1	i	Į	ł	AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII
	1	i .	l			SHYKEEPLTERIKYD
413	1763	A	3361	3	474	PIPVRWNSLEGRLLRGYEQHANDGKDYISRN
	1	1	İ	Ì		*DLRSWTAADMAAQITKRKWEAEEFAEQIKA
	Į .	i	1	ļ		YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI
	1	1	1		İ	PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS
	[-	[			Į.	GHFLPTDRVSYSEAASSDHAQGSDVSLTACK
		1				V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL
	1			1	1	GVHMVDKDTERDIEMKRQLRRLRELHLYST
			ł			WKKYQEAMKTSLGVPQRERDEGSLGKPLCP
	ł	ł	ł	ł	,	PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP
	i		1	İ		EASNQSLLTVAHADAGTQTNGDLEDLEEHGP
,	İ	ļ	1			GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ
	1	1	1	Ì		ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE
		i				EHVPGQTVSEEATGVHMMQVDPATLAKSDL
İ	1	l	ı	1		EDLEEHVPEQTVSEEATGVHMMQVDPATLA
[		1	1		•	KOLEDSTITGSHOOMSASPSSAPAEEATEKTK
1	1			1		VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD
	1	1				IFNIF
415	1765	A	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP
413	1705	1.	1 3303			PSFSPSL
416	1766	A	3373	42	651	ROEKMGLGEIGASGVLRSMLKERKKQNMKG
410	1760	^	3373	122		NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS
		ļ				APGLCSPLHPLQPQQEASTCPSGTLQGREKAA
		i	1			PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD
	1	1	1		1	QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL
		1	1	1	1	RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG
	1	]	ļ			PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2061	EAODPRACGPDAGGRFAARDAPGNSLRPPPS
41/	1707	A	3362	2	2001	SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP
ļ	1	1	ŀ			ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT
·	l	1	1			WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP
				ļ		KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR
	1	1	<b>\</b>		}	ASAGAGAAAALAVGGVRGAGGARGTGGY
1			1	1		GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S
1	1	1	1			· · · · · · · · · · · · · · · · · · ·
1			ſ	1		TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV
		1	1			HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP
1		[				R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP
1		1		1		SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY
		1				PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH
1	1	1		1		SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP
	İ	1		l		PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC
	1			1	l	SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD
1	1	1		1	1	HSSCEGHPDLHAGREMPAAPGLSELERVRFT
				1		VGCGGLASGISSASVSGLSPNRAGGPGQGDW
l		1		1		EMYPVSWQTQESGGQG/SPKTGR*VGMLQA
1				1		GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS
		1	1	1	1	GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS
1					1	Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA
1	ļ	Ì			ŀ	KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	I <sub>A</sub>	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV
1.0	1,00	1.,	1 2370	1		RYGRGKNQKMYEASIKDSDVEGGEVLYLVH
1	1	-		1	İ	YCGWNVRYDEWIKADKIVRPADKNVPKIKH
1						RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN
1	1			1		PPFOTNPISWKWYPKLDLTDAKNSDTAHIKSI
		1		1	1	EITSILNGLQASESSAEDSEQEDERGAQDMDN
1	i	1	1	1		NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE
ì						

SEQ ID NO. of nucl- colide sequence where the color of peptide colide sequence where the color of peptide colide sequence where the color of peptide color of peptide color of peptide sequence where the color of				Long	D. 414	Deadigted and	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence uence unice de la contragonation orrespondis on grant first au fin au de fresidue of peptide residue of peptide sequence p		SEQ ID	Met	SEQ	Predicted	Predicted end	Des A courtie Acid E Glutamic Acid
Sequence   USSN   Olestion   Ol			hod				E-Phenylalonine G-Glycine H-Histidine
Sequence							I-Isolawine K-Lycine I = Ricine
### Briton and residue of peptide residue of peptide sequence peptide sequ	eotide						1 Management National National Proline
### ### ### ### ### ### ### ### ### ##		uence		1			M=Memionine, N=Asparagme, r=rome,
Personal Registry   Pers	uence			914			Q=Giumine, K=Arginite, S=Serice,
			Ì				1=I hreonine, v=valine, w=I ryptopnan,
				1		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
NKYHADLUSKPYSKSPERIKRDEVISKERGKR   YEEDYTKKRKDVKNDTTINSSKRYKKRKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKKEBKAKKKSL   KRYCNTESCL KTGSPCKEBKAKYKGL KEKKK SLRKTGFYSGSFEVAEKRIKLINNSDERL OND   KAKDRKOVWSKIGQOWPKKTL KLEPGDSDTE   AASFPEIPAPEGOVAESLQTVAEESCSFSV   ELEKPFPVNDGKPIESKTYEVNDRKAEFPSS   GNFSA*PIEPAPLIKIRHIOSL ZYGGSRQSS   VTVSEPLAFNQESVESSETDSTTEVDSVAGS   QOKECK	!		l	ì	peptide		
YEEDEVTKRREDVKROTTURSSRPQIKGELEKANKISLCMEN   SSNSSSDEDEEPTKAKNITTURSSRPQIKGELEKANKISLCMEN   SSNSSSDEDEEPTKAKNITTYKKYNGLEERKR   SLRTTGFYSGFSEVABERIKLLINSISERLQNS   RAKDRKDVWSSIQQWPKKTI, KELPEDSDFYE   AAASPPHPAFEGVAEESLCTVAEESCSPTY   AAASPPHPAFEGVAEESLCTVAEESCSPTY   AAASPPHPAFEGVAEESLCTVAEESCSPTY   AAASPPHPAFEGVAEESLCTVAEESCSPTY   KELPPTVNVDKYEEEKTVEVNDRKAEFPSS   GSNFSA*PHPYLLINRILHOSL-QKGSRQOQS   VTVSEPLAPNQEEVSSIKSETDTSTURDSVAGE   LQDLQSEEPLASRIP*QCGLKQ**SAKRTIS* KSLYASEKSERCSGRKFIKKAEKKPSNSGK   QQKECK   QQKE					sequence		nucleotide insertion
RYCNTEECLKTGSPOKKEEKAKNKESLCMES   SINSSSDEDEERIKAKNITTIKKYNGLEEKEK   SINSTGFYSGFSEVAEKRIKLINNSDERI, ON   RAKDRKDWSSIGGOWPKKTILKELFSDSDTE   AAASPPIRPAPEGOVAESLIQTVAEESCSPSV   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVEVNDRAEPPSS   ELEKPPVNVDSKPIEKTVENDRAEPPSS   ELEKPPVNVDSKPIEKTVENDRAEPPSS   ELEKPPVNVDSKPIEKTVENDRAEPPSS   ELEKPPVNVDSKREKTENDSKAEPPSS   ELEKPPVNVDSKPIEKTVENDRAEPPSS   ELEKPPVNVDSKAEPPSSICTORIANS   KSLTASEKSERCSGRRKFIKAEKKPSNSG   QQCECK   QRCCISHIGQGIGDACWELYCLEHGIQP   NGVVLIOTQDQQLENAKMEHINASPTDTFCE   TRAGKHVPRALFVDLEFTVIDGIR   TRAGKHVPRALFVDLEFTVIDGIR   TRAGKHVPRALFVDLEFTVIDGIR   TRAGKHVPRALFVDLEFTVIDGIR   A 3408   IDIO							NKVHADLVISKPVSKSPERLRKDIEVLSEDTD
SSNSSSDEDEETKAKMTITKKYNGLERKK   SLRTTGFYSGFSVAKRRILLNISDERLQNS   RANDRIDVWSSIQGQWPKKTLKELFSDSTTE   AAASPHEAPEGVAESELQTVAEESCSPSY   ELEKPPYNIVDSKPIEEKTVEVNDRKAEHPSS   GSNFSA*IPLPYLH.NRLHOSL*OKGSRQOSS   VTVSPLAFNQEEVRSIKSETDSTTEVDSVAGE   LQDLQSERF*LASR*PCQCELKQ**SARTRTS*   KSLYRSEKSEEGSGRKVFIKAEKK*SNSGTK   KSLYRSEKSEEGSGRKVFIKAEKK*SNSGTK   KSLYRSEKSEEGSGRKVFIKAEKK*SNSGTK   KSLYRSEKSEEGSGRKVFIKAEKK*SNSGTK   KSLYRSEKSEEGGSGRKFIKAEKK*SNSGTK   KSLYRSEKSEEGGSGRKFIKAEKK*SNSGTK   KSLYRSEKSEEGGSGRKFIKAEKK*SNSGTK   KSLYRSEKSEEGGSGRKFIKAEKK*SNSGTK   KSLYRSEKSEEGGSRKFIKAEKK*SNSGTK   KSLYRSEKSEEGGSRKFIKAEKK*SNSGTK   KSLYRSEKSEEGGSRKFIKAEKK*SNSGTTK   KSLYRSEKSEEGGSRKFIKAEKK*SNSGTTK   KSLYRSEKSEEGGSRKFIKAEKK*SNSGTTFCC   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TRAGKHVPRALFVDLEPT/UGARVQWRDLGSPPLP   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVQMSDLAKAELV*P   TAGKHVPRALFVDLEPT/UGARVAELV*P   TAGKHVPRAL	i		ļ.		1		YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR
SIRTIGFYSGFSEVAEKRIKLINNSDELQNS   RAKDRKDVWSSIQQWPKKIKELFSDSDTE   AMASPPHAPEGVAEESIQTVAEESCSFSV   ELEKPPPVNTDSKPIEEKTVENDKAEFFSS   GSNFSA*IPLPYLHINKIHOSI*QKGSRQQSS   VTVSEPLAPNQEEVRSKSETDSTIEDSVAGE   LQDQSERF*IASRF*QCQELKQ**SARTRIS*   KSLYRSEKSERGSGRKFIKKAEKK*\$NSOK   QQKEGK   QRECISHIQQAGIQIGDACWELYCLEHGIQP   NGVVLDTQQDQLENAKMEHTNASPDTFCE   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKH			ļ		ļ	1	RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
SIRTIGFYSGFSEVAEKRIKLINNSDELQNS   RAKDRKDVWSSIQQWPKKIKELFSDSDTE   AMASPPHAPEGVAEESIQTVAEESCSFSV   ELEKPPPVNTDSKPIEEKTVENDKAEFFSS   GSNFSA*IPLPYLHINKIHOSI*QKGSRQQSS   VTVSEPLAPNQEEVRSKSETDSTIEDSVAGE   LQDQSERF*IASRF*QCQELKQ**SARTRIS*   KSLYRSEKSERGSGRKFIKKAEKK*\$NSOK   QQKEGK   QRECISHIQQAGIQIGDACWELYCLEHGIQP   NGVVLDTQQDQLENAKMEHTNASPDTFCE   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLEPTVIDGIR   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKHYPRALFVDLATAGRA   TRAGKH			ļ		l		SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK
RAKDRKDVWSSIQGGWPKKTLKELFSSDTE	i i			1	1		SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS
AAASPPHPAPEGOVAEESIGTVAEESISPSY			1			ł	RAKDRKDVWSSIOGOWPKKTLKELFSDSDTE
ELEKPPYNVDSKPIECKTVEVNDRAEFFES   GSNFSA*IPJYHLINRALIFOSL*QKGSRQOS    VTVSEPLAPNQEEVPSIKSETDSTUGSUGS    LQDLQSERP*LASRP*GCCLKQ**SARTRTS*   KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK    QQKECLSPSWDVARPPYVDFPYVFVVLV    QWKHEHVQQAGLELTSGDLPALASQSARIT    GVNHCAQPRGIFFI    A 3409 355 1326 ADSNLESCWQELGIGPWGGDWRVEQVGAS    ASLRFPREVCSIRLITAVSLLSIFJSAWLGL    LYLVSPLENEPREMITLISEYHERVASQQQL    QQLQAELDKLHKEVSTVRAANSERVAKLVF    QRENEPFYKRPYALSSAMILQKTSHDV    ADRITAYPWNRTSPWNYARPPTVLEPHVFP    GNCWAFEGDQGQVVIQLPGRQUSDITLQHFP    PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ    VYDETEVSLCKFTFDVEKSEQITFHQNDPPA    AFPKVKIQLISNWGHPFTCLYRVRAHGVRT    SEGAEGSAQGFH    APFKVKIQLISNWGHPFTCLYRVRAHGVRT    SEGAEGSAQGFH    A 3412 2 421 EFDAQPSIGALVYKRP*ATTGSDPGPKRGMN    YLVSCSMRSPPSCKGEFGTARDYTHGVFF    SEGAEGSAQGFH    QARSPPGGWLGSATRVRRFHNHP/RGH/HSP    VDTAGAPASPOPDVCE    QARSPPGGWLGSATRVRRFHNHP/RGH/HSP    VDTAGAPASPOPDVCE    A 3420 91 706 DAQRAHYSSVGPASULQKQQDGAVKESGR    RGGVRSFSRAAAAMAPIKVGDAPAVEYFEG    QARSPPGGWLGSATRVRRFHNHP/RGH/HSP    VDTAGAPASPOPDVCE    XFHLJGFVEQAEALKAKGVQVVACLSVNDA    FERKRYVQVERBUTPGAPULRVYADKDKKG    VTVGVLDNDAPIFYSTFQATVLSDRNG    SRAVYHFSMSONARGOFYLDAGTGALDVV   SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    SFLDYETTKEYTLRVRAQDGGRPYLDAGTGALDVV    QTTSONTRNRFSTISTGSGGGGLVSLAPLDYKLE    QTTSONTRNRFSTISTGSGGGGLVSLAPLDY	i i		l	1			AAASPPHPAPEEGVAEESLOTVAEEESCSPSV
GSNFSA*PILPYLHJRLHQSL*QKGSRQQSS	{		l l		1	ì	ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS
1769	) }		ł	1 .	1	<b>}</b>	GSNESA*IPLPYLHLNRLHOSL*OKGSROOSS
LQDLQSERP1ASRR*GQCELKQ**SARTRIS*   KSLYRSEKSERCSGRKFIKKAEKKP*SNSGK   QQKEGK     419   1769   A   3399   206   463   QRECLSIHIGQAGIQIGDACWELYCLEHGIQP     420   1770   A   3408   1010   685   RRLSFP*IWSSVLVTQARVQWRDLGSPQPLP     421   1771   A   3409   355   1326   RRLSFP*IWSSVLVTQARVQWRDLGSPQPLP     421   1771   A   3409   355   1326   ADSNLIESCWGELGGPWGGDWRVEQVGAS     422   1772   A   3412   2   421   ADSNLIESCWGELGGPWGGDWRVSQVQVG     424   1774   A   3412   2   421   ADSNLIESCWGELKFIFDVEKSEQTFHLQNDPPA     425   1773   A   3420   91   706   APPLY REPROBLEMENT     426   1774   A   3421   4   7688   RQVTRVFRPAJSVRQRQAL     427   1774   A   3421   4   7688   RQVTRVFREPALAWING REPROBLEMENT     428   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     429   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     420   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     421   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     422   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     424   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     425   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     426   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     427   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     428   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     429   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     420   1774   A   3421   4   7688   RQVTRVFREPALAWING REPVENDENT     421   1774   A   3421   4   7688   RQVTRVFRENT     422   1774   A   3421   4   7688   RQVTRVFRENT     423   1774   A   3421   4   7688   RQVTRVFRENT     424   1774   A   3421   4   7688   RQVTRVFRENT     425   1774   A   3421   4   7688   RQVTRVFRENT     426   1774   A   3421   4   7688   RQVTRVFRENT     427   1774   A   3421   4   7688   RQVTRVFRENT     428   1774   A   3421   4   7688   RQVTRVFRENT     429   1774   A   3421   4   7688   RQVTRVFRENT     420   1774   A   3421   4   7688   RQVTRVFRENT     421   1774   A   342	]		1	ļ	ļ		VTVSEPI APNOFEVRSIKSETDSTIEVDSVAGE
	! !					1	LODLOSERE*LASRE*COCELKO**SARTRTS*
1769	1 1		1		1	1	LODEOSEKE FASIG COCKERVA FIKK PINSGK
1769		l	1	1	1	1	
NGVVLDTQQDQLENAKMEHTNASFDTFCE					1		QQKEGK
TRAGKHYPRALFYDLEPT/YDGIR	419	1769	Α	3399	206	463	QRECLSIHIGQAGIQIGDACWEL1CLEHGIQF
1770		<u> </u>	1	1			NGVVLDTQQDQLENAKMEHINASFDIFFCE
PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV	j	]	1	ľ		1	TRAGKHVPRALFVDLEPTVIDGIR
PGFKRFSCLSLPSSWDYRHFSPRPVNF/HYFLV VMGFHHVGQACIELLTSGDLPALASQSARIT GVNHCAQPRGHFH  421 1771 A 3409 355 1326 ADSNLESSCWQELGLGPWGGDWRVEQVGAS ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGI LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKLHKEVTVRAANSERVAKLVF QRLMEDFYRKPYVALASSVGASIDLQKTSHDY ADRNTAYFWNRFSFWNYARPFIVILEPHVFF GNCWAFEGDQGQVVQLDFBRYQLSDITLQHD PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFTFDVEKSEQTFHLQNDPPA AFPRVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPEGKGEPGTARDYTPMGRPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RG/HSP VDTAGAPASPGPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQDGAVKESGR/ CARSPPGWLGSAT/RVRRPHNHP/RG/HSP VDTAGAPASPGPDVCE  424 1774 A 3421 4 7688 RGVYRSFRAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA PVTGEWGRARKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAFNISQL  424 1774 A 3421 4 7688 RQVYRNGTRAVISVEDDNDNAPQ FSEKRYVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYFITKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFSTFQATVLSVFEGVEAR DHGTPALTASASVSVTALDVNDNNFTFQPF FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASSVSVTALDVNDNNFTFTQPF TYTVRLNEDAAVGTSVVTVSAAVDRDAHSSVITY QITSONTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVVTDANTH RPVFGOSSHYTVNVNDERPAGTTVVVLISATDE	420	1770	A	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP
421 1771 A 3409 355 1326 ADSNLIESCWQELGLGPWGGDWRVEQVGAS ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKIKEYSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRITAYFWNRSSYMYARPFTVLEFHTPV GNCWAFEGDQGQVVIQLPGRVQLSDITLQHF PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFFFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEFGTARDYTPMGRPPP PPPSVSPGPLFGSLALAFHSPEPHFWEQQPPRG QARSPPGGWLGSATRVRRPHNHP/RGH/HSP VDTAGAPASPGPVCE  423 1773 A 3420 91 706 DAQRATYSSVGPAVSLRQRQQDGAVESGR/ RGGVRSFSRAAAMAPIKVGDAPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTFGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGONRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL GVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGGFYLDAQTGALDVV SPLDVETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDNDNAPIFVSTFPQATVLESVPLGY LVHVQADAAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTTTQPE YTVRLNEDAAAGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYYLAVTASDGTGOTAQTTVVVLISATDE			1		1	į	PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
421 1771 A 3409 355 1326 ADSNLIESCWQELGLGPWGGDWRVEQVGAS ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKIKEYSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRITAYFWNRSSYMYARPFTVLEFHTPV GNCWAFEGDQGQVVIQLPGRVQLSDITLQHF PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFFFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEFGTARDYTPMGRPPP PPPSVSPGPLFGSLALAFHSPEPHFWEQQPPRG QARSPPGGWLGSATRVRRPHNHP/RGH/HSP VDTAGAPASPGPVCE  423 1773 A 3420 91 706 DAQRATYSSVGPAVSLRQRQQDGAVESGR/ RGGVRSFSRAAAMAPIKVGDAPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTFGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGONRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL GVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGGFYLDAQTGALDVV SPLDVETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDNDNAPIFVSTFPQATVLESVPLGY LVHVQADAAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTTTQPE YTVRLNEDAAAGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYYLAVTASDGTGOTAQTTVVVLISATDE	į į	1.		i .		1	VMGFHHVGQAGLELLTSGDLPALASQSARIT
ASLRPPREVCSIRPLFTAVSILLSLFLSAFWLGIL LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKHKEVSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNFSFWNYARPPTVILEPHVFP GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFFTPVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGFKRGMN YLVSCSMRSPESGKGEFGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGWLGSATRVRRPHNHP/RGH/HSP VDTAGAPASPOPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGRKVNLALFKGKKGVLJCFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLYEFGNRFLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVRIGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVRIGSTTAAVFLSVEDDNDNAPQ SNAVHYSIMSGNARGGPYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDNDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTAASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSOSGGGLVSLAIPLDYKLE RQYVLAVTASDGTRQDTAQVVNVTDANTH RPVFOSSHYTYNVNEDRPAGTTVVLISATDE			1		ł		GVNHCAOPRGHFH
ASLRPPREVCSIRPLFTAVSILLSLFLSAFWLGIL LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKHKEVSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNFSFWNYARPPTVILEPHVFP GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFFTPVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGFKRGMN YLVSCSMRSPESGKGEFGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGWLGSATRVRRPHNHP/RGH/HSP VDTAGAPASPOPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGRKVNLALFKGKKGVLJCFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLYEFGNRFLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVRIGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVRIGSTTAAVFLSVEDDNDNAPQ SNAVHYSIMSGNARGGPYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDNDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTAASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSOSGGGLVSLAIPLDYKLE RQYVLAVTASDGTRQDTAQVVNVTDANTH RPVFOSSHYTYNVNEDRPAGTTVVLISATDE	421	1271	<del>                                     </del>	3400	355	1326	ADSNLIESCWOELGLGPWGGDWRVEQVGAS
LYLVSPLENEPKEMITLSEYHERVRSQGQL QQLQAELDKLHKEVSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNRFSFWNYARPTVILEPHVFF GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTIGGANSAPRDFAVFFLLSFTTHQGLQ VYDETEVSLGKFTFDVEKSBIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVYFKRP*ATTGSDPGFKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPPDVCE  423 1773 A 3420 91 706 DAQRATYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTFGCS KTHLPGFVEQABALKAKGVQVVACLSVNDA FYTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAFNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVELSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGGFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFTSTPQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPITFTQPE YTVRLNEDBAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTIDANTH RPVFOSSHYTVNNVEDRPAGTTVVLISATDE	421	17/1	^	3403	1 333	1320	ASI REPREVESIRELETAVSLLSLELSAFWLGL
QQLQAELDKLHKEVSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRITATYEVNRFSFWNYARPPTVILEPHVPF GNCWAFEGDQGQVVIQLFGRVQLSDITLQHP PFSVBHTGGANSARPDFAVFFLLSFTTHQGLQ VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH 422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGFKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLFGSLAIAPHSPEPHPWEQQPPRG QARSPPGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE 423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EFGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL 424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTFPQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRTSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNNVEDRPAGTTVVLISATDE	1	ļ	1		1		1 VI VSPLENEPKEMI.TLSEYHERVRSOGOOL
ORLNEDFYRKPDYALSSVGASIDLQKTSHDY			1	1	<b>,</b>		OOL OAEL DEL HEEVSTURAANSERVAKI VE
ADRNTAYFWNRPSFWNYARPPIVILEPHVFP GNCWAFEGDQGQVVIQLPGRVQLSDITLQHFP PFSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESCKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPVCE  423 1773 A 3420 91 706 DAQRAIYSSVGAVSLRQRQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLLADPTGAFGKET OLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVQVREDVITGAPVLRVTASDRDKG SNAVVHYSIMSGNARGOFYLDAQTGALDVV SPLDYETTKEVTLRVRAQDGGRPPLSNVSGL VTQVULDNDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH	l	1	1		Ĭ		OPLATED EVEN PROVAL SSAGASIDI OKTSHDA
GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPPGTGLTCSLAPNIISQL AVVEYSIMSGNARGGFYLDAVTASDRDKG SNAVVHYSIMSGNARGGFYLDAVTGALDVV SPLDYETTKEYTLRVRAQDGGRPLSNVSGL VTVQVLDINDNAPIFVSTFPQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	1				1		QKLNEDY VKKYD I KLSS V GASIDLQKI SIID I
PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFTFDVEKSEIQTFHLLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDCTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGGFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGGNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYYLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	1	l	1		1		
VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQDDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGRNRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	ļ	İ		1	1		GNCWAFEGDQGQVVIQLFGKVQLSDI1LQFI
AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH  422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNNVEDRPAGTTVVLISATDE	ļ	1	1				PPSVEHIGGANSAPROFAVFFLLSFFIHQULQ
422 1772 A 3412 2 421 EFDAQPSIGALVYFKP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  423 1773 A 3420 91 706 DAQRATYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL LNVEPDGTGLTCSLAPNIISQL SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVILHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDANTH RQYVLAVTASDGTRQDTAQIVVNVTDANTH RQYVLAVTASDGTRQDTAQIVVNVTDANTH RQYVLAVTASDGTRQDTAQIVVNVTDANTH RQYVLAVTASDGTRQDTAQIVVNVTDANTH	1	1		1			VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA
422 1772 A 3412 2 421 EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPUVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPPCTGLTCSLAPNIISQL PVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPPCTGLTCSLAPNIISQL SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTFPQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFSITSQSGGLVSLALPLDYKLI RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPYFOSSHYTVNVNEDRPAGTTVVLISATDE		1		1	1		
YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGGFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	ļ						SEGAEGSAQGPH
PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPSPGPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGOFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFFSITSQSGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	422	1772	A	3412	2	421	
QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE  423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE		1	1		j		YLVSCSMRSPESGKGEPGTARDYTPMGRPPP
423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE							PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG
423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE	•	1	1			ļ	QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
423 1773 A 3420 91 706 DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE		1			İ		VDTAGAPASPGPDVCE
RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE	122	1773	A	3420	91	706	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE	423	1//3	Δ.	3720	1.		RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG
KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGOFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE	-	Ι'			1		EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNNNEDRPAGTTVVLISATDE	1	1		}	1 .	(	KTHLPGFVEOAEALKAKGVOVVACLSVNDA
DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNFFSITSQSGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNNVEDRPAGTTVVLISATDE				1	1		FVTGEWGRAHKAEGKVRI LADPTGAFGKET
LNVEPDGTGLTCSLAPNIISQL  424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDNDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE		1	1		l		DI LI DDSI VSIEGNERI KRESMVVODGIVKA
424 1774 A 3421 4 7688 RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGGFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE		1	1	ŀ	1	1	I ARTEDOCTOL TOSI ADMISON
FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTFFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE				1			LINYELDO I GLI COLLI AMENO GIRCODADALA DO
SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE	424	1774	A	3421	4	7688	KQVIKVGIKVLOSIIAAVILSVEDDINDINAPQ
SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE		1	1				FSEKRYVVQVREDVIYGAPVLRVIASDRDKG
VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE		1	- (		1		SNAVVHYSIMSGNARGQFYLDAQIGALDVV
LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE		1	ľ		1		SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL
FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE		1			1		VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE		1	1	-	1		LVLHVQAIDADAGDNARLEYRLAGVGHDFP
DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	1		.		1		FTINNGTGWISVAAELDREEVDFYSFGVEAR
YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE	1		1	1	1		DHGTPALTASASVSVTALDVNDNNPTFTQPE
QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVIDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE			Ī	1			YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY
RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFOSSHYTVNVNEDRPAGTTVVLISATDE		1	1	1	1		OITSGNTRNRFSITSOSGGGLVSLALPLDYKLE
RPVFOSSHYTVNVNEDRPAGTTVVLISATDE		-	1		ļ		POVU AVTASDOTRODTADIVVNVTDANTH
DTGENARITYEMEDGENDADTGAVITOA		[				1	PRIECELLALIVITATION DE LE LA LINE DE LE LA LINE DE LE LA LINE DE LA
TOTAL TOTAL STATE OF THE PROPERTY OF THE PROPE	1		1			1	TACEMADALA A MEDICACATA A MISALDE
PUDLISHING THE PROPERTY ET		1				1	DIGENAKII I PMEDSIPQPKIDADIGAVI I QA
ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI	1						
LVNDVNDNAPQFLRDSYQGSVYEDVPFF1SV				İ			ELDY EDQVSY I LATTAKDNOT POSTO
LQISATDRDSGLNGRVFYTFQGGDDGDGDFI							LVNDVNDNAPOFLRDSYQGSVYEDVPPFTSV

PCT/US01/03800

	OTO YES	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID					D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	•	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ucilco	Ì	<b>!</b>	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ĺ		l .	i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ĺ	1		sequence	/=possible nucleotide deletion, \=possible
1	ł	1	l	peptide		
1		i	L	sequence		nucleotide insertion
		Τ –	· .			VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
		1	1		ŀ	GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
	l	ł	l	Ī	1 .	VFVEENSPIGLAVARVTATDPDEGTNAQIMY
		1	1	ì	}	QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
		1		Ì		YVLVIOATSAPLVSRATVHVRLLDRNDNPPV
ļ	1	1	1	ļ		LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
	1 .	ł		ł	l	
1	1	1	ł			DISDSLTYSFERGNELSLVLLNASTGELKLSR
1	ŀ		1.		1	ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
ļ	1	1	1	Į.	1	TIITDEMLTHSITLRLEDMSPERFLSPLLGLFIQ
1	1	l .	1	1	1	AVAATLATPPDHVVVFNVQRDTDAPGGHILN
	1	1	1	1	1	VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
1	1		1			LLTAISAORVLPFDDNICLREPCENYMRCVSV
	1	1	1	1	ł	LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
1 .	1	1	1	1		TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
	1	1	1	1		
	}	1		1		CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
		Į.		1	1	CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
1	1	-	į.			AHSFITFRGLRQRFHFTLALSFATKERDGLLL
l l	ł		j	1	1	YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
ı	1	1	ĺ	1		TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
1	1		1	į.		QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
	1		1	1	İ	VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
	1	i	}	1		VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
i	ĺ	1	1			ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
	1	1	1	ľ	i	VNQWDAFSCECPLGFGGKSCAQEMANPQHF
		1	1	1	1	LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
	1	1		1	1	
ĺ	1	1		1		GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
		1		ì		QASSLRLEPGRANDGDWHHAQLALGAIGGP
	1		1			GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
i	j	]				GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
			1			SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
				1		CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
			ł	Ì		EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
ì	)	1	į.	}		RIDOPCPRGWWGHPTCGPCNCDVSKGFDPDC
			ļ			NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
ì	1	1				
	ı		1	1		SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
,			1	1		AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG
1			1	1		LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
	1	1	1			NCTSITFSELKGFAERLQRNESGLDSGRSQQL
1	. [			1		ALLLRNATQHTAGYFGSDVKVAYQLATRLL
		l	1	1	, .	AHESTQRGFGLSATQDVHFTENLLRVGSALL
1	.	Ī	1	1		DTANKRHWELIQQTEGGTAWLLQHYEAYAS
1	1		1	1	1	ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
1	1	1	ł			FAGAKLPRYEALRGEQPPDLETTVILPESVFR
		1		1		ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
	1			1		GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
	1	1		[	1	DEPARTMENT OF THE PROPERTY OF
	1	1			1	RPIINTPVVSISVHDDEELLPRALDKPVTVQFR
				1	1	LLETEERTKPICVFWNHSILVSGTGGWSARGC
			1			EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
Ì						NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
'	ĺ	-				RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
1	ļ	- [	1		1 .	DLPFACTVIAILLHFLYLCTFSWALLEALHLY
1		1				RALTEVRDVNTGPMRFYYMLGWGVPAFITG
ĺ	1					LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
1	1	ĺ				VAFAVSMSVFLYILAARASCAAQRQGFEKKG
1			1	1		PVSGLOPSFAVLLLLSATWLLALLSVNSDTLL
1	1	J				
1		1		1	1	FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
	1		1	1		LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
, ,		1	1			RLYQP\YGDSAGSLHSTSRSGKSQPSYTPFLLR
				1	1	EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ
	ì			1	1	H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

PCT/US01/03800

SeQ ID NO: of mucle of peptide sequence						D. Paris	Amino acid sequence (A=Alanine C=Cysteine,
muchoted confide sequence contraspond years of the contraspond of the contraspond of the contraspond of the contraspond of the contraspond of the contraspond of the contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide sequence contraspond of peptide contraspond of	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A-Alamic C-Cysteme,
uence op	NO: of	NO: of	hod				D=Aspartic Acid, E=Oludinic Acid,
uence    09496   914	nucl-	peptide					F=Pnenylalanine, G=Glycine, H=filstidine,
uence  9446 914 914 914 914 914 914 914 914 914 914	eotide	seq-	ļ	USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
uenoe    14	1		1	09/496	correspondi	to last amino	
mino seid residue of peptide sequence peptide sequence peptide sequence peptide sequence of peptide sequen			İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
residue of peptide sequence    V=Tyrosinic, X=Unknown, *=Step codon, /-possible nucleotide insertion	uchoo	ĺ	1	1		of peptide	T=Threonine, V=Valine, W=Tryptophan,
peptide		1	ł	ļ			Y=Tyrosine, X=Unknown, *=Stop codon,
Sequence   micleoide insertion	ļ			1		sequence	/=nossible nucleotide deletion. \=nossible
### EPEPEERAAFPGEGGWISSLIOPGARENJPHS ### TPEDGGPOFGKAPWFODFGTTARESSGNGAP ### EPERLRENGDALSREGSLGPLPGSSAOPHKOIL KKKCLPTISERSSLILLPLEGCTGSSRGSNGAP #### EPERLRENGDALSREGSLGPLPGSSAOPHKOIL KKKCLPTISERSSLILLPLEGCTGSSRGSNAE GTVDEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH ### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDSGSEFLFNFNLH #### GTPUEDGARRATFALISSYPOPPPSLPA AGTAVKRGLQPOGAGATSTGTGAATGGL GPAWAAPSAVGPCCCPSISTTPSGMRSARP SLGCLPSWASVGTHEPPQPQGPSPDLCSVV KREFQGGPWAGMVILLRISAADPARAPGPDS NAGNAAPSAVGPCCCCPSISTTPSGMRSARP SLGCLPSWASVGTHEPPQPQGPSPDLCSVV KREFQGGPWAGMVILLRISAADPARAPGPDS NAGNAAPSAVGPCCCCPSISTTPSGMRSARP SLQSALQOPATCCSEPSLSOVELRHLANGSYS OAGHSVASVASVASVASVASVASVASVASVASVASVASVASVAS	t	ŀ	ı	1		ŀ	
TREDGGPGFGKAPWPGDFGTTAKESSGNGAP EERLRENGDALSREGGSGPLGPSSAPKIGDI KKKCLPTISEKSSILALPLEOCTOSSRGSSASE GSGGGPFSPPPRGLEQCLAKOVMPLAMSIKA GTVDEDSGSEFLFNFLH  GEPAVOSCDGCGTGRSCPSCWLLVAPGLLSSSSS REASSVEAEDAPLQPASIFPVSQGSRGPEGSL GSAGCLPGPSGCPGCTQSSCPWLLVAPGLLSSSSS REASSVEAEDAPLQPASIFPVSQGSRGPEGSL GSAGCLPGPSLATALSISPVPGPPSLPA AGTAVKRGLQPP GAGATSTFGTGAATGGL CGPAWAAPSAVEPCCCEPSISTTSGMRSAR SLGCLPSWASVGTEHPPGPQGPGPS-DLCSVA KREFQRGWAGAWILHRISAADPRAAGPDS NLQSALQQPATGGSPAAVYSPFIGLWGA*PH EYG*PGHSLQPTAPADR*PAGCPSSTDLSWAFARCPSS NLQSALQQPATGGSPAAVYSPFIGLWGA*PH EYG*PGHSLQPTAPADR*PAGCPSSTDLAMA*FGSSS NLQSALQQPATGGSPAAVYSPFIGLWGA*PH EYG*PGHSLQPTAPADR*PAGCPSSTDQWLAHLRIGEKE EVTVGSLKTSAVPSTSTMSQBFELLISGMEKP FLPKTTDS F		<u> </u>	<u> </u>		sequence		
### FERLERENGDALSREGSLEPLFGSSAQPHKOLL KKKCLPTISKESSLEPLEPLGTGSSRGSASE GSRGGPPSRPPRQSLQEQLNGVMPIAMSIKA GTVDEDSSGSEFLFFNFLH  ### GEPAVQSCOGGTQRSCPWLLVAPGLLSSSS RASVSRADAPLOPASHIPVQGQSRGPEGSL GSRCLPGDPLGARRATRAHSFVPGPPSLDA AGTAVKBGLQPG-GA/GATSTPGTGAATGGL CGPAWAAPSAVGPCCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGL CGPAWAAPSAVGPCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGL CGPAWAAPSAVGPCCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGL CGPAWAAPSAVGPCCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGL CGPAWAAPSAVGPCCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGL CGPAWAAPSAVGPCCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGL CGPAWAAPSAVGPCCCCFSISTTPSQMRSAR SLGCLSPWASPGTGAATGGAATGACCL KREETORD ALL QARACTAL COLOR SLGCLSPWASPGTCLSPTATGGAATGGL CGPAWAAPSAVGPCCCCPSISTTPSQMRSAR SLGCLSPWASPGTCLSPTATGGLWAPP SLGCLSPWASPGTALDAPAGRACHAGCC CGPAWAAPSAVGPCCCCPSISTTPSQMRSAR SLGCLSPWASPGTCLSPTATGGLWAPP SLGCLSPWASPGTALDAPAGRACHAGCC CGPAWAAPSAVGPCCCCPSISTTPSQMRSAGC CGPSCLSPATSTSTSGLWAPP AGGAGACTATATATATATATATATATATATATATATATAT		]	i e	•		<b>l</b> .	EEEEEEEAAFFGEGGWDSLLGFGAEALI LIIS
KKKCLPTISEKSSLIR.PLE.QCTGSSRGSSASSASSASSASSASSASSASSERS GROGPSRSPPRSEJ QQQUQMYDIAMSIKA GTVDEDSSGSEFLFNFLH  425 1775 A 3429 155 1417 GEPAVQSCOGGTORSCPWILVAPGLISSSSS RAASVREAEDAFLQPASHFVSQGSRGFGSSSSSSSASSASSASSASSASSASSASSASSASSASSAS	ł	}	l	į .		ĺ	TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
GSRGGPPSRPPRQSLQSQLNGVMPIAMSIKA GTVDEDSSGSEFLFFNFH.H  425 1775 A 3429 155 1417 GEPAVQSCOGGTQRSCPWLLVAPGLILSSSS GASACLPGDPLGARRATRAHSPVPGPPSLAVA AGTAVSRGLQPG-QAVGATSTFGTGAATGGL CGPAWAAPSAVGPCCCESISTTPSOMRSARP SLGCLPSWASPGTGAATGGLACTGAATGGL CCPAWAAPSAVGPCCCESISTTPSOMRSARP SLGCLPSWASPGTAPADR*PAGINDRVSNSI YELLENGQRAGTCVLEYATPLQTLPAMSQYS QAGRSRGBRIEQAKLFCRLEDLLADAPESQN NCRLIAYQPADDSSPSLSQSVLRH.RQERE EVTVGSLKTSAVPSTSTSMQGPELLISOMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWISWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSPRQPMHYLQALRAVGGICQDYDV SVGESGAGGNRQGLAQRIPQLLPSDKRFP AFGCARIPPNFEVG*MRGKEGDGGRVSQA KAGPHCSRLALTOSHDANDFTASNGDPR SSQSLHCLSPRQPMHYLQALRAVGGICQDYDV SVGESGAGGNRQGLAQRIPQLLPSDKRFP AFGCARIPPNFEVG*MRGKEGDGGRVSQA KAGPHCSRLALTOSHDANDFDFPDECEGK RGDFHLPRLPADTLHTGAQTPLPRAQLPYPST HRPVFPEISIVASYRRCLPQIGVADDS DMRLLDGDDGFLCPRGVPAARDIVQFVPFR DFKDVSPGFFELLSDSSASHPRSDRLPPD VLLRTREPSWPP*STTSPSDDPASPTLPLTPNHI TVPTLAARSALAKCVLAKVPRQVVEYYASQ GLGRGVQOGQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	ì	i	İ				EERLRENGDALSREGSLGPLPGSSAQPHKGIL
425 1775 A 3429 155 1417 GEPAVQSCDGGTQTQSCPWILVAPGILISSSSS RAASVREAEDAFLQPASIBFVSQQSQGGGGSQCASCASCASCASCASCASCASCASCASCASCASCASCASC	ł	ł				l .	KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
425 1775 A 3429 155 1417 GEPAVQSCDGGTQTQSCPWILVAPGILISSSSS RAASVREAEDAFLQPASIBFVSQQSQGGGGSQCASCASCASCASCASCASCASCASCASCASCASCASCASC	Į.		ł	1	Į.		GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
RAASVREAEDAPLQPASHFPVSQCSRGFGSL GSAECLIOPLGARRATRAHSPYGFPSLPA AGTAVKRGLQPG*GAGATSTFGTGAATGGL CQPAWAPSAVQPCCCCPSISTTPSQMRSARP SLGCLPSWASVPGTEHPPGPQGPGPS*DLCSV* KREFQRGPWAGNVTLHRISANDRAPGPDS NLQSALQQPATGCSPDAAVYSFPIGLWGA*PP EYG*PQRISLPG*TAPADR*PAGIKDRVYSNSI YELLENGGRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLPCRTLEDILADAPESON NCRLIAVQEPADDSSFSLSQGVLRHLRQEBKE EVTVSSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPIPTQVAIDFTASNGDPR SQSGLHCLSFRQPNHYLQALRAVGGICQDYDY SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNEPVG*MGKBCARGGGRVSQAE KAGPHCSRLAATGSSIDPAINFOPENPECEGK RGDPHLPRLAPADLTHTGAGTPAALDVPSTS HPRPVFTEISGVIASYRRCLPGIQLYGPTNVAP IINRVAPPAQRZQSTGQATXYSVLLVLTDGV VSDMAETRTAVRASRLPMSII/VGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPR DFKDVSPFGPFRIKDSSASHPFKSDLRLPFD VLLRTREPSWPPSTSPSDDPASFTLPTPHII TVPTTLAAFSALAKCVLAEVPRQVVEYYASQ GISPQAPPLGTLATTPSPSP  6CGSCWPAWPTLRREGPASAGARLGRKAPW GLGRGVQDGRPLRFCFVLRPRAPFIAPVLSGA ASRPASGDCTAAKCVLAEVPRQVVEYYASQ GISPQAPPLTLATTPSPSP  7 9748 GCGSCWPAWPTLRREGPASAGARLGRKAPW GLGRGVQDGQQQQQQQQQQQQQQPPP PPPPPPPQLPQAPPLPPPPPPPPPP	1	1	1	1		1 •	GTVDEDSSGSEFLFFNFLH
RAASVREAEDAPLQPASHFPVSQCSRGFGSL GSAECLIOPLGARRATRAHSPYGFPSLPA AGTAVKRGLQPG*GAGATSTFGTGAATGGL CQPAWAPSAVQPCCCCPSISTTPSQMRSARP SLGCLPSWASVPGTEHPPGPQGPGPS*DLCSV* KREFQRGPWAGNVTLHRISANDRAPGPDS NLQSALQQPATGCSPDAAVYSFPIGLWGA*PP EYG*PQRISLPG*TAPADR*PAGIKDRVYSNSI YELLENGGRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLPCRTLEDILADAPESON NCRLIAVQEPADDSSFSLSQGVLRHLRQEBKE EVTVSSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPIPTQVAIDFTASNGDPR SQSGLHCLSFRQPNHYLQALRAVGGICQDYDY SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNEPVG*MGKBCARGGGRVSQAE KAGPHCSRLAATGSSIDPAINFOPENPECEGK RGDPHLPRLAPADLTHTGAGTPAALDVPSTS HPRPVFTEISGVIASYRRCLPGIQLYGPTNVAP IINRVAPPAQRZQSTGQATXYSVLLVLTDGV VSDMAETRTAVRASRLPMSII/VGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPR DFKDVSPFGPFRIKDSSASHPFKSDLRLPFD VLLRTREPSWPPSTSPSDDPASFTLPTPHII TVPTTLAAFSALAKCVLAEVPRQVVEYYASQ GISPQAPPLGTLATTPSPSP  6CGSCWPAWPTLRREGPASAGARLGRKAPW GLGRGVQDGRPLRFCFVLRPRAPFIAPVLSGA ASRPASGDCTAAKCVLAEVPRQVVEYYASQ GISPQAPPLTLATTPSPSP  7 9748 GCGSCWPAWPTLRREGPASAGARLGRKAPW GLGRGVQDGQQQQQQQQQQQQQQPPP PPPPPPPQLPQAPPLPPPPPPPPPP	105	1275	+	2420	155	1417	GEPAVOSCDCGCTORSCPWLLVAPGLLSSSSS
GSAECLFODPLOGARRATRAHSPYOFOPPSIPA AGTAVKRGLOGYGAJGAGTSTGTGAATIGGL CCPPAWAAPSAVOPCCCCFSISTTPSQMSARRS SIGCLPSWASPGTEHPOPGOFORS*DLCSV* KREFQRGPWAGMVILHRISAADPARAPGPS NLQSALQPATGCSEPAAVYSPFIGLWGA**P EYG*PQHSLFG*TAPADR*PAGBCRVYSNSI YELLENGGRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLFCKTLEGHADAPESQN NCRLIAYQFADDSSFSLSQEVLRHLRGEKE EVTVGSLKTSAVSPSTIMSQPSPLLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 ALWILSWLOPDLLIPTOVAIDFTASMGDPR SSQSLHCLISFRQPRIPTIONAP SSQSLHCLISFRQPRIPTIONAP SSQSLHCLISFRQPNIPYLQALRAVGGICQDYD/SVGSSGSAGGNRQGGLAQRIPQLFLIPSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYG*MKGKEDGGRVSQAE KAGPHCSRLALTGSHIDPANFDFENDECLIFSDKRFP AFGFGARIPNFEYGMKGKGDGGSTVSQAE KAGPHCSRLATTGVSSASSAFPKSDLRIJPPDH VILLTREPSSTRALTGVSATSULLTDGVVFP DFKDVSPFGFFRLGAGARGTAMATLELAKAFESL AFKKDRVSHLCLIC GENGAGSKGTAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	425	17/3	) ^	3429	133	1417	PAASVPEAEDAPLOPASTHPVSOGSRGPEGSL
AGTAVKRGLQRG*GAGATSTPGTGATGGL CGPAWARPSAVPCCCCESISTTPSQMRSARP SLGCLPSWASVGTEHPPGPQGPGPS*DLCSV* KREPGRGPWAGWILHRISADRAPGPDS NLQSALQQPATGCSEPAAVYSEPIGLWGA*P EYG*PQHSLPTAPADR*PAGKIDRVYSNSI YELLENGQRAGTCVLEYAIFLQTLPAMSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEQAKLPCKTLEDILADAFSQVS QAGFSEDRLEPTPQVALDETPTQVALDFTASNGDPR SSQSLHCLSFRQPNHYLQALARAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNEVG*MRGKGARAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNEVG*MRGKGAGGRVSQAE KAGPHCSRLALTGSSIDPANHPDFENPECEGK RGDFHLPRLAPAQLPYPST HPPVYTRISGVMASYRRCLPQIQLYQFVFR DFKDVSPFGFFRKDSXASHPFKSDLRLPFFD VLLRTREPSWPP*STFSSDDPASFTLLTPNHI TVYTLAAFSALAKCVLAEVPRQVVEYYASQ GISPQAFPRCTLATTPSPSP  QCSCWFAWFRLRRGPASAGARLGRKAFW GLGRGVQDGRPLRFCFYLRPRAPFIAPVLSGA ASPPASGDCAAKCVLAEVPRQVVEYYASQ GISPQAFPLTLATTPSPSP  GOSCWFAWFRLRRGPASAGARLGRKAFW GLGRGVQDGQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	1	i	1		ļ		COARCI DODDI CARP ATP AHSPVPGPPPSI PA
CGPAWAPSAVGPCCCCFSISTTPSQMRSARP SLGCLPSWASSPGTEHPGPGQFGPS*DLCSV* KREFCRGPWAGMVILHRISADPARAPGPS NLQSALQQPATGCSPPAAVSPFIGLWGA*P EYG*PQHSLPG*TAFADR*PAGIKDRVYSNSI YSLLENGQRAGTCVLEYATPLQTLFAMSQVS QAGFSREDRLEQAKLPCRTLEDILADAPESQN NCRLIAVQEPADDSSFSLSQEVLRI-LRQEBKE EVTVGSLKTSAVSPSTIMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSFRQPNIHYLQALRAVGGICQDYD/ SVESGAGGNRGGLAGRIPCHLPSDKRPP AFGFGAIPPNETVG*WRGKEGDGRYXQAE KAGPHCSRLAITGSHDFAHDQFLFLLPSDKRPP AFGFGAIPPNETVG*WRGKEGDGRYXQAE KAGPHCSRLAITGSHDFAHDQFLFLLPSDKRPP AFGFGAIPPNETVG*WRGKEGDGRYXQAE KAGPHCSRLAITGSHDFAHDQFLFLLPSDKRPP AFFRYFIEISGVIASYYRRCLPQQLYLPTFNAP INRVAPPAQRSGTGQATKYSVLLVLTDGV VSDMASTRTATVRASRLPMSIIIVGVGHADPS DMRLDGDDGPLRCPGGVPAARDIVQFVPFR DFKDVSPPGFPRLKDSSASHPFKSDLRLPPFD VLRTREPSWPFSTFSSDDASFTLPFTDHT TVPTILAAPSALAKCVLAEVPRQVVEYYASQ GISPQAPPCTLATTPSSPS GQSCWPAWPRLRRRGPASAGARLEKAPV GLIFGRVQDGRPLRCFCFVLRPRAFTIAPVLSGA ASRPEASGDCARGETAMATLEKLMKAFESL CSGQQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQCPPPPPPPPPP	l			1		1	A CONTRACT ODCACA (CATCEDCTGA ATGGI
SIGGLPSWASDGTEHPFGQGGGFS-DLCSV* KREFGQEPWAGMVLHRISAADPARAFGPDS NLOSALQQPATGGSEPAAVYSPPIGLWGA**P EYG*PQHSLPG*TAPADR*PPAGKDRYYSNSI YELLENGQRAGTCVLEYATFLQTLFAMSQYS QAGFSREDRLEQAKLFCRTLEDILAD#ESQN NCRLIAYQEPADDSSTSLAGEVLRHLRQEBKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWULSWLQHDLLPTFTQVAIDFTASNGDPR SQSJLKCLSRQPNHYLQALRAVGGICQDYD/SVGESGAGGNRQGLAQRIPQLFLLPSDKRFP AFGFGARIPNFEVG*MRGKEGGRYSQAE KAGPHCSRLALTGSSHDFANNDEENPECEGK RGDFHLPRLPADTLHTGAQTPLFRAQLPVPST HPRYVFEISGVASYRRCIQULYGFTNVAP IINRYAEPAQREQSTGQATKYSVLVLTDGV VSDMAETRTAAVRASRLEMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSAFPKSDLRJPFD VLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTILAAPSALAKCVLAEVPRQVVEYYASQ GCGSCWPAWPBLRTRGPASAGARLGRKAPW FORDATATAVASRLEMSIIIVGVGNADFS DMRLLDGDGPLRCPRGVPAARDIVGPVPFR DFKDVSPPGPFRLLOTSSHPKSDLRLPFD VLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTILAAPSALAKCVLAEVPRQVVEYYASQ GCGSCWPAWPBLRTRGPASAGARLGRKAPW GCGSCGCOQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	1	1	1	1			AGTAVKRGLQPG*GA/GATSTPGTGAATGGL
KREFGRGPWAGMVILHRISAADPARAFGPDS NLOSALQOPATGCSEPAAVSPIGLWGA*P EYG*PQHSLPG*TAGADR*PYAGIKDRVYSNSI YELLENGQRAGTCVLEYATPLQTLAMSQYS QAGFSREDRLEQAKLFCRILEDILADAPESQN NCRLIAYQEPADDSSTSLSQEVLRRILNQEEKE EVYTGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRITDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SQSLHCLSPRQPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLISDKERP AFGFGARIPPNFEVG*MRGKGGDGGRVSQAE KAGPHCSRLALTGSHDFADNFDPENECGK RGPHLPRLPADTLHTGAQTPLFRAQLPVPST HPRPVFNEISGVIASYTRCLPQIQLYGPTNVAP HINRVAEPAQREQSTGGATKSVLLUTDGV VSDMAETRTAIVRASRLPMSIHVGVGNADFS DMRLLDGDDGFLRCPRGVPAARDIVG*VPFS DFKDVSPPGPFRLDSSSAFPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPILPTPHTH TVPTILAAPSALAKCVLASSHIPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPILPTPHTH TVPTILAAPSALAKCVLASKIPKSVLLVTDGV VSDMAETRTAIVRASRLPMSHVGVGNADFS DMRLLDGDDGFLRCPRGVPAARDIVG*VPFS DFKDVSPPGPFRLDSSSAFPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPILPTPHTH TVPTILAAPSALAKCVLAKKIKATPSLS GCQSCWPAWFRLRRERGPSAGARLGRKAPW GLQRGVQDQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	1	1		1	1		
NLQSÁLQQPATGCSEPAVYSPPIGLWGA**P EYG*PCHSLPG*TAPADRAYNAGIKDRYYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLFCRTLEDILADAPESON NCRLIAYQEPADDSSPSI_SQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SQSLHCLSRQPMHYLQALRAYGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRPP AFGGARIPPNFEVG*MRGEGGGRYSQAE KAGPHCSRLALTGSHDFANFDPENPECEK RGDFHLPRLPADTLHTGAQTPLFRAQLPVPST HPRPVPTEISGVASFRCQLYGYTMVAP INRWAPPAQREQSTGQATKYSVLLYLTDGV VSDMAETRTAIVRASRLFMSILIVGVGNADFS DMRLLDGDDGFLRCPRGVPAARDIVGVPVPR DFKDVSPPGPFRLDSSASIPPKSDLRLPPD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVEYYASQ GCQSCWPAWPELTRRGPASAGARLGRKAFW GCGSCWPAWPELTRRGPASAGARLGRKAFW GCGSCWPAWPELTRRGPASAGARLGRKAFW ASRFASGDCRAGRETAMATLEKLMKAFESL KSCQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ		1	1	Ī	[	1	
NLQSÁLQQPATGCSEPAVYSPPIGLWGA**P EYG*PCHSLPG*TAPADRAYNAGIKDRYYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLFCRTLEDILADAPESON NCRLIAYQEPADDSSPSI_SQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SQSLHCLSRQPMHYLQALRAYGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRPP AFGGARIPPNFEVG*MRGEGGGRYSQAE KAGPHCSRLALTGSHDFANFDPENPECEK RGDFHLPRLPADTLHTGAQTPLFRAQLPVPST HPRPVPTEISGVASFRCQLYGYTMVAP INRWAPPAQREQSTGQATKYSVLLYLTDGV VSDMAETRTAIVRASRLFMSILIVGVGNADFS DMRLLDGDDGFLRCPRGVPAARDIVGVPVPR DFKDVSPPGPFRLDSSASIPPKSDLRLPPD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVEYYASQ GCQSCWPAWPELTRRGPASAGARLGRKAFW GCGSCWPAWPELTRRGPASAGARLGRKAFW GCGSCWPAWPELTRRGPASAGARLGRKAFW ASRFASGDCRAGRETAMATLEKLMKAFESL KSCQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ				١ ،	1	1	KREFQRGPWAGMVILHRISAADPARAPGPDS
### FYG** PORISIPG**TAPADR**PAGIRDRY*SNS)	}	1	1	1	1		NLOSALOOPATGCSEPAAVYSPPIGLWGA**P
YELLENOGRAGTCVLEYATPLOJITAMSCYS QAGFSREDRILEQAKLFCRTLEDILADAPESON NCRLIAYQEPADDSSFSLSQEVLRHRQEEKE EVTVGSLKTSAVPSTSTIMSQEPELLISGMEKP LPLRTIDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTTQVAIDFTASNGDPR SSOSLHCLSPROPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRPQLFLPSDKRPP AFGGARIPPNFEVG*MRGKGDGGGRVSQAE KAGPHCSRLALTGSHDFANHFDFENFECEGK RGDFHLPRLPADTLHTGAQTPLPAQLEVPST HPRPVFHEISGVIASYRCLPQQLQQLYQTPNVAP INRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGFFRLKDSSASHPPKSDLRLPFPDI VLLRTREPSWPP*STFSSDDPASPTLPLTNHH TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGARPCTLATTPSSSDDPASPTLPLTNHH TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGARPCTLATTPSSSDPASPTLPLTNHH TVPTLAAPSALAKCVLAEVRQVVEYYASQ GISPGARPCTLATTPSFSDDPASPTLPLTNHH TVPTLAAPSALAKCVLAEVRQVOVEYYASQ GISPGARPCTLATTPSFSDDPASPTLPLTNHH TVPTLAAPSALAKCVLAEVRQVOVEYYASQ GISPGARPCTLATTPSFSDDPASPTLPLTNHH TVPTLAAPSALAKCVLAEVRQVEYYASQ GLPGRVQDGRQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	1	!	1	1	1	1	EYG*POHSLPG*TAPADR*P\AGIKDRVYSNSI
QAGFSREDRLEQAKLFCRTLEDILADAPESQN NCRLIAYQEPADDSSPSLSQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  A 3431 1662 369 AIWWLSWLQHDLLFTPTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGCQDYD/ SSQSLHCLSPRQPNHYLQALRAVGGCQDYD/ SVGESGAGGNRQGGLAQRIPQLFLPSDKRFP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDFANHFDESPECGK RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRRVFHEISGVIASYRRCLPQIQLYGFTNVAY INRVAEPAQREQSTGAYKSYLLVLTDGV VSDMAETRTATVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIVGVGNADFS DMRLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIVGVGNADFS DMRLLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIVGVGNADFS DMRLLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIVGVGNADFS DMRLLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIVGVGNADFS DMRLLDGDDGPLRCFGVAXSVLVLVLTDGV VSDMAETRTATVRASRLPMSIIVGVGNADFS DMRLLDGDDGPLRCFGTVLRVASVQ GISPGAPRCTLATTPSSSP  GCQSCWPAVPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRRAFTAPVLSGA ASRPASADCDCAGRETAMATLEKIMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPP	1	j	1	1	1	1	VELLENGOR AGTCVLEY ATPLOTLEAMSOYS
NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE BYTVGSIKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS  426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SQSLHCLSFRQPMYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRPP AAFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDPANTPDFENPECEGK RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HERPVFNEISGVIASYRRCLPQIQLYGFTNVAP INRVAPPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASTRRCLPQIQLYGFTNVAP INRVAPPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASTRRCLPQIQLYGFTNVAP INRVAPPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASTRRCHPSIIVCGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPFD VLRTREPSWPP*SPTSPSDDPASFTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTSSPSP GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPP PPPPPP		1	1 .	1	l .		OAGESPEDRI FOAKI FORTI EDILADAPESON
### BUTVGSILKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS    A	1	1		ļ	1	1	MCDI IN NOEDY DOGGEGI GUENI BHI BUEEKE
426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQFLFLLPSDKRPP AAGFGARIPPNFEVG*MRGKGDGGRVSQAE KAGPHCSRLALTG\SHDFAINFDPENPECEGK RGDFHLPRLPADTLHTGAGTPLPRAQLPVPST HERPVFIDEISGVIASSYRCCLPQIQL\YGFTNVAP ID\RVAPPAQREQ\STGQAT\YSVLLVLTDGV VSDMAETRTAT\VRASRLPMS\IIU\YGV\SHAPPSD D\RKLDGDDGPLRCPRGVPAARDIV\QFVPFR DF\KDV\SPFGPFRL\KDS\SAS\IPP\KSDLRLPPFD VLLRTREPS\WPP*STPSPDDPLS\SPTLLTIP\HIT TVPTL\AAPSALAKCVL\AEVPRQVVEYYASQ GISPGAPR\CTLATTPS\SPS GC\SCCVPA\WFRLATT\RSF\RT\ABTA\LEL\HK\AF\SEL\ SS\QQQ\QQQQQQQQQQQQQQQQQQQQQQ\QPPP PPPPPPQL\PQPPPQ\QPLL\PQ\PPPPPPPPP G\RAVABEPL\HR\K\EL\SAT\KDR\WH\CL\LS SS\QQQQQQQQQQQQQQQQQQQQQQQQQQQ\QPPP PPPPPPQL\PQPPPQ\QPL\LQ\LD\\PQ\PPPPPPPPPP	}	1	Ì				DETERMINATION AND CONTRACT ACCEPTED TO COME AND CONTRACT AND CONTRACT ACCEPTED TO COME AND CONTRACT ACCEPTED TO COME AND CONTRACT ACCEPTED TO COME AND CONTRACT ACCEPTED TO COME AND CONTRACT ACCEPTED TO COME AND CONTRACT ACCEPTED TO COME AND CONTRACT ACCEPTED TO
426 1776 A 3431 1662 369 AIWWLSWLQHDLLPTFTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGICQDYD'S VGESGAGGNNQGLAQRIPQLFLPSDKRFP AFGGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDFANDFDEPBECGK RGDFHLPRLPADTLHTGAQTPLPRAQLEVPST HFRPVFTEISGVIASYRGLPQIQLYGPTNVAP INRVARPAQREGSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMELLDGDDGPLFQRGVPAARDIVQFVPFR DFKDVSPPGPFPRGVPAARDIVQFVPFR DFKDVSPPGPFRLADSSASHPPKSDLRLPPFD VLLRTREPSWPP*9FTSPSDDFASPTLPLTPNHI TVPTILAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	İ	1	1	,		-	
SSQSLHCLSPRQPNIHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLIPSDKRPP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSEIDPANDFDENPECGK RGDFHLPRLPADTAHTHOADITPLPRAQLPVPST HPRPVFTIEISGVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTATVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVFFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLVAAPSALAKCVLAEVPRQVVEYYASQ GISPOAPRPCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRTCFYTURPAPFIAPVLSGA ASRPEASGDCTALATPSPSP — PPPPPPPQLQPPPQAQQQQQQQQQQQQQPPPP PPPPPPPQLQPPPQAQPLLPQPQPPPPPPPPPP	1	· F		1	Ì		LPLRTDFS
SSQSLHCLSPRQPNIFYLQALRAVGGICQDYD/ SVGESCAGGNRQGGLAQRIPQL-FLLPSDKRPP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDFANIPDENPECEGK RGDFHLPRLPADTAHTHAQITPLEPAQLPVPST HPRPVFNEISGVIASYRRCLPQIQLYQFTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTATVASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVFR DFKDVSPPQPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLVAAPSALAKCVLAAEVPRQVVEYYASQ GISPOAPRPCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRTCFYLRPRAPFIAPVLSQA ASRPFASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPPP PPPPPPPPPQLPQPPPPAQPPPPPPPPPP	426	1776	A	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR
AFGFGARIPPNFEUG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDFANFDENPECGK RGDFHLPRLFADTLHTGAQTPLPRAQLPVPST HPRPVFTIEISGVIASYRRCLPQIQLYQFITNVAP IINRVAEPAGRSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLVAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	1	1		1	1	1	SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/
AFGFGARIPPNFEUG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDFANFDENPECGK RGDFHLPRLFADTLHTGAQTPLPRAQLPVPST HPRPVFTIEISGVIASYRRCLPQIQLYQFITNVAP IINRVAEPAGRSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLVAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	Ì	ł		1	İ		SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP
KAGPHCSRLALTGSHDFAINFDPENPECEGK RGDFHLPRLPADTILHTGAQTPLPRAQLPVPST HPRPVTFIEISGVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIVGVGRAADFS DMRLLDGDDFLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRCTLATTEPSPS GCQSCWPAWPTPRSPSPS GCQSCWPAWPTPRSPSPS GCQSCWPAWPTRRGGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKIMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILLSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSETSVLLNVLLQQ KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVTOALELLQQLFRTPPPEL LQTITAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSKQKGKVLLGEEBALEDDS ESRSDVSSALTASVKDEISGELAASSGVSTPG SAGHDITTGQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	1	ì	1	1	İ	1	AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE
RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRPVFIEISQVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTATVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASIPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKIMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLINKVIKALMDSNLPRLQELL YKEIKKNGAPRSLRALMWRAELAHLVRPQK CRPYLVNLLPCLTRTSKRPESSVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPII RRTAAGSASVCQHSRRTQYFYSWLINVLIG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTREMEVSPSAEQLVQVYEL TLHHTQHQDHNVYTGALELLQQLFRTPPEL LQTILTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRCKKGKKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVKLLSASFLLTGGKNVLVPDRDVRV	ì		1	ŀ	1		KAGPHCSRLALTG\SHDFAINFDPENPECEGK
HPRPVFIEISGVIASYRRCLPOIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFFLKDSSASIPPKSDLRLPPFD VILBTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTIJAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPPCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL  — KSFQQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPOAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAGVSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLLGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEBALEDDS ESRSDVSSAALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVKLLSASFLLTGGKNVLVPDRDVRV	1	1	1		1	1	RGDFHI PRI PADTI HTGAOTPLPRAOLPVPST
IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASIPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLJAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPPCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLERRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAFFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAPPPPPPPPPP		1				1	UDD DVENEIGGVIA SVERCI POIOI VGPTNVAP
VSDMAETRTAIVRASRLPMSIIIVGVCNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLMAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPHHRPKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFGKLIGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KMASFGNFANDNEIKVILKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTILLIGVLLILTRYLVPLLQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDENVLVPDRDVRV	İ	Į		1	ļ	1	
DMRLLDGDDGPLRCPRGVPAARDIVQFVFFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPAPRCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPP PPPPPPPPQLPPPQLPQPPPPPPPP	· ·	1	1	Į.	ì		
DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VILRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLVAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQPPPQPPPQAQPLLPQPCPPPPPPPPPPPPPPP	1	l.		ĺ	Ĭ.	}	VSDMAEIKIAIVKASKLIMSIIIVO VOIVADIS
VLLRTREPSWPP*SPTSPSDDPASPTI.PLTPNHI TVPTILAAPSALAKCVLAEVPRQVEYYASQ GISPGAPPPCTI.ATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRI.RRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLQCPLPOPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTILLILGVLLTLRYLVPLLQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITGQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQAHLLKNMSHCRQPSDSSVDKP VLRDEATEPGDQENKPCRKGDIGQSTDDDS APLVHCVVRLLSASFILLTGGKNVLVPDRDVRV	1	i		1	1	1	
TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPPRCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQPPP PPPPPPQLPQPPPQAQPLLPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLIL.GVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTILTAVGIGQLTAAKEESGGRSRSGIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTILQADSVDLASCDLTSS ATDGDEEDILSHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQAHLLKNMSHCRQPSDSSVVKF VLRDEATEPGDQENKPCRKGDIGQSTDDDS APLVHCVVRLLSASFILLTGGKNVLVPDRDVRV	Ì	l l	I	i			DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD
GISPGAPRPCTLATTPSPSP  427 1777 A 3446 79 9748 GCQSCWPAWPRILRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPDLPQPPPOAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VIRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	ì	ļ	1	1		1	VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI
427 1777 A 3446 79 9748 GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	·	1 .	1		1		
GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPP	-	1		i			
GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPP	427	1777	A	3446	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQPPPP PPPPPPPPPPPPPPPP	727	1	'-		1		GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
- KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPOLPQPPPQAPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLIL.GVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			ĺ		1		ASRPEASGDCRAGRETAMATLEKLMKAFESL
PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			i		1		KSF00000000000000000000000000PPPP
GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV						_	DDDDDDDU DUDDDDU VUDI I DUDUDDDDDDDDDDDD
ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANIKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			-		1	1	
ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			1		1		
YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTOHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	1		1		1	1	ENIVAÇOVENOPETÇELLGIAMELITLEGODDA
CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLIL GVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	1				Į.		ESDVRMVADECLNKVIKALMDSNLPRLQLEL
KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTTKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV					l	.	YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			1	1		[	CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			1		1	1	KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV		}	i		1	1	RRTAAGSAVSICOHSRRTOYFYSWLLNVLLG
KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			Į.	1	1	i	LLVPVEDEHSTILII.GVI LTLRYLVPLLOOOV
TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	1				{	1	KDTSI KOSEGVTRKEMEVSPSAFOI VOVVEI
LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV				1	1		TI INTOLODIDANTATO A DI LOOI EDTEDDEI
AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	'		-	1	1		LOT TAYOUTOUT A AVERGO CHERCOTTELL
ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQVSAVPSDPSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			-		l l		LUILIAVGGIGULIAAKEESGGKSKSGSIVELI
SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV							
ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV			1		}		
ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV		-	- [		Į.		SAGHDIITEQPRSQHTLQADSVDLASCDLTSS
TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV		1	1	1	1		
GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	1		1	1	1		
FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV		1	1	1	1		
VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV	1	1			1		ERNSSMAL OO AHLI KNMSHCROPSDSSVDKR
APLVHCVRLLSASFLLTGGKNVLVPDRDVRV				1	1		ALEDEVIED COUNTRICK CONCUCTODOR
			1		1	1	A DE TRECOME E CA CEL E TOCCASTA COMPENSOR
SVKALALSCVGAAVALHPESFFSKLYKVPLD			1	1	1	1	
		l i					SVKALALSUVGAAVALHPESFFSKLIKVPLD

[ ATA TA	000 10	1 1 1 1 1	GEC.	Donalis 4 - 3	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted		
NO: of	NO: of	hod .	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1 .	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uciice	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		1	914			
ì	ł	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	1		peptide		/=possible nucleotide deletion, \=possible
				sequence	l	nucleotide insertion
<b></b>	<del> </del>	<del></del> -	<del> </del>	boquision		TTEYPEEQYVSDILNYIDHGDPQVRGATAILC
	1	1	1		i	
1	Ì	l	ł			GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
1.						ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
İ	1	ļ		ļ	l	SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
		1				LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
		ľ				LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
	{		i			1
1		]	}	'	ì	IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
		1			1	KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
1		1	1			VTMENNLSRVIAAVSHELITSTTRALTFGCCE
1		1	1	1	1	ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
1		1		1		KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
1			Į		1	ILAGNLLAASAPKSLRSSWASEEEANPAATK
I .	1		1		I .	
1		1	[	1	, 1	QEEVWPALGDRALVPMVEQLFSHLLKVINIC
1	1	1	1		1	AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
		1	}			GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
1.		1	1	1.	İ	GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
	Ì		1	1		NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
1				1		
ł	1	ł	ł	ł	1	ATLQDIGKCVEEILGYLKSCFSREPMMATVC
		Į.	ì		1	VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
	Ĭ		1		· ·	QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
		1	1	i		SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
1 .		ł	1	1		NLTSVTKNRADKNAIHNHIRLFEPLVIKALKO
1	1	1	1	1	[	1
1	1	1	1	1		YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
1	i			i		DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
1 .		Į.	1		1	FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
·		i		1	1	KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
1	1	1				LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
ļ	1	1	1	1	ŀ	CHKENEDKWKRLSRQIADIILPMLAKQQMHI
İ		i	1	1	}	
1			i	1	1	DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF
1		1	1			VTPNTMASVSTVQLWISGILAILRVLISQSTED
1	l	1	i		1	IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
1		1	ļ		İ	EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
1	}		j		j	KOLKVEMSEOOHTFYCOELGTLLMCLIHIFKS
1 .	1	1				
1					1	GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
	1	1	1		}	ARSMITTHPALVLLWCQILLLVNHTDYRWW
1		1	1			AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
1		1	1	1	I	AKLGMCNREIVRRGALILFCDYVCQNLHDSE
1		1	1	1	1	HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
1		{ ·	1	1		
		i	1	1	1	AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
1	l	1	1	1		HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
1	i	1	1			ACRRVEMLLAANLQSSMAQLPMEELNRIQEY
1	1	1				LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
1			1			PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
1	Į.	1	1	1	į.	SOCWTRSDSALLEGAELVNRIPAEDMNAFM
1	1	1	1	1	1	
	ļ	1		1		MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
1	I		I	1		AREVTLARVSGTVQQLPAVHHVFQPELPAEP
}	1	ì	1	ļ		AAYWSKLNDLFGDAALYQSLPTLARALAQY
1	ı	1	1	[	1	LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
1	1		1	1		SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
1	1		1	1		
į.	1			1	1	WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
i		1	1	l		EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI
1		1	1	1		TAACEMVAEMVESLQSVLALGHKRNSGVPA
ŀ	1	1	1	1	1	FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
- 1	1	1		i		
1		1	1	1		WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
1	ļ	]		1,		INTLGWTSRTQFEETWATLLGVLVTQPLVME
1	1	1	1	1	1	QEESPPEEDTERTQINVLAVQAITSLVLSAMT
ı	1		1	1		VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
1	1	1		j	1	LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
i	1		1	1		PVPSLSPATTGALISHEKLLLQINPERELGSMS
1	1	1	1	1	Į.	
					<u></u>	YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of, peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV KLSVDRVNVHSPHRAMAALGLMLTCMYTG KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL
					·	FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
429	1779	A	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKIGTKQGAVRARKECRGEMA SGETDSE
430	1780	A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE ASKCNIVCTQPRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEVHER\SVQS DFLLIILKEILQKRSDLHLILMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIEETGFVLE KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPQLQVISNAMNLLRKIGACELNEPK LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTIYNAYLGWKKARQEGGYRSEI TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA GLYDNVGKIIYTKSVDVTEKLACIVETAQK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY LRETTLITFPPVLLFGGDIEVQHRERLLSIDGW NYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK MSLENDKILQIITELIKTENN
431	1781	A	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV

SEQ ID Mot of nucleotide peptide peptide per control of peptide per	000 ID	SEO ID	Mak	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide sequence coliders of the contraspondity of the contrasp		•					D-Amortic Acid E=Glutemic Acid
Sequence			noa	1			E-Phenylolonine G-Glycine H-Histidine
1782							Information V-I wine I = I strine
1782						corresponding	
### ### ### ### ### ### ### ### ### ##	seq-	uence					
Presidue of peptide   Pe	uence	İ		914			
#possible nucleotide deletion, "possible nucleotide insertion ### squence ###			1	1	1		
1782		l .		į	residue of	sequence	
432   1782	ĺ			į	peptide		/=possible nucleotide deletion, \=possible
1782   A 3478   416   23   QLRRLITPIRKTY/YSS*IBLAWH**RMMQID   QWFRRESPEIDLCK YSS*ISDREAKAR/WE   CSLFNRWCYKNWM/HYOKKRI*VQTLHPS   QKLKSK WIEDLNYECRITKLIDQEYFGILGY   SRAINSGSR   CLAPSOPPERNGMQPLILLIPPLLYQQLLHS   SRAINSGSR   CLAPSOPPERNGMQPLILLIPPLLYQQLLHS   SLAGAPGESTLLVRTSKLLVGLIGLICLVWILLI   QTRSILALQHILTSSAPLIAAFIAACSCSRCS   APRSRCVARPAARTGLPTPAPASSPAPAASPA   PAASPAPAESTAVQPLILIPPRIPASSPAPAASPA   PAASPAPAESTAVQPLILIPPRIPASSPAPAASPA   AASPSPAASSPAPAASPAPASSPARASSPAPASS		l		1	sequence	Ì	
0	432	1782	A	3478		23	OLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
CSLFNKWCYNNWLHVOKKRI VQTLHPS	732	1,02	1	1 2 . , ,	'		OWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE
	l	•	1	1	[		CSI FNKWC/VKNWM/LHVOKKRI*VOTI.HPS
1783   A   3504   1876   552   CLAPCESCOPEKIGMOPILLLLEPILLYQULHS	}			ļ	}	}	ON KICKMINDI MVECBITKI I DOEVPGDI GV
1783   A   3504   1876   552	i			1		ł	
SIGAPGESTILVRTSKILVGI.GLQLLVWILLI  (TRSIALALQH.ITSSAPILLASPATVACSGERCS APRSRCVARPARTGI.FTPAPASSPAPAASSPAPASSPATASSPAPAASSPATASSPAPA  ASPAPAASSPATASSPAPASSPATASSPAPATASSPAPASSPAPASSPATASSPAPA	<b></b>		L			<u> </u>	SKALINGUSK
OTTSILALQHITISSAPLIAAPTIAVCSCRSC APPSRCVARPAASIPAASPAASPAASPAASPAPAASPAARGIQHTAASSPAASPAASPAPAASPAASPAARGIQHTAASSPAASPPAASPPAASPPAASPPAASPPAASPPAAS	433	1783	A	3504	1876	552	
APRSECVARPAARTIGLITIFAPASSPAPAASSPA PAASPAPAESTAQOPLILITYAPPAASPPLAASPPL PAASPPAASTAQOPLILITYAPPIPAASPPLAASPPLAASPPL AASPPAASPPLASPVITASPPLPAASPVITASPPLY ASPPVHASPPVHASPPVHASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVASPPVH	t		l .				
PAASPAPAESTAVQQLILLEKPPPAAGSPPRA GAPPPRAASPALTASPPLP AASPSPAASPAPAASPVLTASPPLPAASPPA ASPAPAASPVLTASPPLPAASPPA ASPAPAASPVLTASPPLPAASPPAH ASPSPHASPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVA		1	l		į		
GAPPRPAASPEASPUTASPPLPAASPUTASPPLPAASPSPAASPUTASPPLPAASPSPAASPVATASPPLPAASPASPUTASPPLPAASPASPUTASPPLPAASPASPASPUTASPPLPAASPASPATASPPASPVHVASPPVHANDASPPVHASPPVHASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHVASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHVASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATONASPPVHAATO			l	Į.	ļ	1	
AASPSPAASPVLTASPPLASSPSLA ASPAPAASPVLTASPPLASSPALAASPVIT ASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVA				1.	Į.	i	PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP
AASPSPAASPVLTASPPLASSPSLA ASPAPAASPVLTASPPLASSPALAASPVIT ASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVASPPVHTASPPVHVA	1	1	l			<b>.</b>	GAPPPRPAASPSPAASPAPPAASPVLTASPPLP
ASPPAMASPVITASPPLAASPPALASPVHT ASPPVHVASPPVHTASPPVHVASPPVHVASPPVHVASPPVHTASPPVHVASPPVSCGGGDSTSDCFPP QPGAVPHELAPSLGGWSHLVAALP  434 1784 A 3516 142 590 GGVNRPRSETEQVKTPVLISSWDVRIPPRPA  435 1785 A 3529 1 3161 MSLVRAALEALDELDLFGVKGGVGCVGVWQWHNLGGLQPLSLEDRLSPGVLGCSALCRSGVWNGVHANGSLQPLSLEDRLSPGVLGCSALCRSGVWNGVHANGLLYPECGSALCRSGVWNGVHANGSLQPLSLEDRLSPGVHCAGSALCRSGVWNGVHANGSLQPLSEDLSPGVHCAGSALCRSGVWNGVGCVGCVGCCRFLRAGGVKGCCRFLRAGGVKGCCRFLRAGGVAGCCRFLRAGGVKGCCRFTAGGVKGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCRFTAGGCCCCRFTAGGCCCRFTAGGCCCCRFTAGGCCCCRFTAGGCCCCRFTAGGCCCCRFTAGGCCCCRFTAGGCCCCRFTAGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		ĺ	1			İ	AASPSPAASPAPPAASPVLTASPPLPAASPSPA
ASPYWIASPPVHTASPPHVHASPPVHTASPPHVHASPPVHVASVQ GRIVERPRAS AS 3516 142 590 GGVNRPSLESEGVKPVLGCSALCRSOV RIKFGRMYTISERSTVILFEGSALCSVOV WHNLGSLQPLSLESVOV GRIVERGLASPVHVAD DEVQHCQSILNSLLPRASTSKEVDASLLSVVS PPAFAVEDSOLVELTKQGRYGGCGR LRDGYKTPKEDPNLYY/ENPAELKLFENIEC EWPL/FWTYFILDGVPSGGAREQVQCKRALEA VLIKGKRIGVPLLPPLSVYSPDRVDEFYQNPHT VDRVPMGKLPHMVGQSLVILGSLMAEGFILA PGEIDPLNRRSTVPKPVQVYSPDEPVGLVSKALEA VLIKGKRIGVPLLPPLSVLAGEN GREIPFLASPPRASLASPP			}		}		ASPAPPAASPVLTASPPLPAASPALAASPVHT
HAVASPPVHTASPPVHVASPPVSCSGDSTSDCTPP  QPGAVPPHSLAPSLGGWSHLVAALP  434 1784 A 3516 142 590 GGVNRPRSETEQVKTPVLJSSWDYRHPPRPA  435 1785 A 3529 1 3161 MSLVRAALEALDELDLFQKVGGPVHVLA  WINLGSLQPLSLEDRLSPQKGCSALCRSGV  WINLGSLQPLSLEDRLSPGVSUHVLA  MSLVRAALEALDELDLFQKVGGGPSVHVLA  DEVQHCQSLINSLLPRASTSKEVDASLLSVVS  FPAFAVEDSQLVELTKQEITKLQGRYGCCRF  LRDGYKTPKEDPPNLTYVFDALEKLFENIEC  EWPLFWTYFILDQVSGGAACQVQEVKEALEA  VLIKGKNGVPLLPELYSVPPDRVDEEVYPPSLEHLEFNIEC  EWPLFWTYFILDQVSGGAACQVQEVKEALEA  VLIKGKNGVPLLPELYSVPPDRVDEEVYPPSLEHGE  EWPLFWTYFILDQVSGGAACQVQEVKEALEA  VLIKGKNGVPLTAEVSTRACGGPVAHMGVLGTSK  LYDIRKTIFTETPGFDQQGPTLALDKKMIVE  MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE  DGTSLNSSLLAALRKMQDEVTGGARVQTGKL  SEFLITSCCTIHLSFMDFGFEGKL YSEDYDDN  YDYLESGNWMNDYDSTSHARGGDEVARVL  DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT  TCDLMSLVTKAKELHVQNVHMYLFTKLFQA  SRSFSHLLDSHPRQCENQVSVRVEHLPRDQ  SGEVDFKALVLQLKETSSLQEADILYMLYT  MKGPDWNTELTYRERSATVRELLTELTGKVG  EIRHWGLIRYISGIRKKVEALDEACTDLLSH  QKHLTVGLPFPEREKTISHCSABEATGEL  MMLSPSAMKNLLHHILSGKEFGVERSVRPTD  SNVSPAISHEIGAVQATKTERTGIMQLKSEIK  QSPGTSMTPSGSSFPSAYDQQSSKDSRQQQW  QRRRRLDGALNRVPVGFYQKVWLQKCH  GLSVEGFVLPSSTTERMFPGEIKFSVHVESVL  NRVPQPEYRQLLVEALLVLTMLADEHNISIGS  INVERVIHANDLFLEGGKTLGADDTMLAKD  PASGICTLLYDSAPSGRFGTMTYLSKAAATY  VQEFLPHSICAMQ  QFELTBELEVKKAEVLQDSLDGRYSTPSSCL							
ASPPVHTASPPVHVASPPHVASPPHVASPPPHVASPPPHVASPPPHVASPPPHVASPPPHVASPPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVHVASPPVSCGDSTSDCFPP QPGAVFPHSLAPSLGGWSHLVAALP GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA SFFVFLV*TGFTALARMVLISWPDLPTSASQ SAGTIGVRHARALLYFEQESHSVTQAGWVQ WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV RTKFGINMVTSRERGTTRLPKEG  ### A 3529 1 3161 MSLVRAALEADEELDLFGVKGGPQSVHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS PPARAVEDSQLVELITKQGITKLQGRYGCCRF LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PPARAVEDSQLVELITKQGITKLQGRYGCCRF LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PPARAVEDSQLVELITKQGITKLQGRYGCCRF LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PPARAVEDSQLVELITKQGTKLQGCTR LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PPARAVEDSQLVELITKQGTKLQGCTR LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PPARAVEDSQLVELITKLQGRYGCCRF LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PPARAVEDSQLVELITALGGTTRLTKQCCCRF LRDGYKTPKEDPNRLYYEPPARVEDASLLSVVS PRAPAVEDSQLVELITALGGTTRLTKQCCCRF LRDGYKTPKEDSQLVELITALGGCCRF LRDGYKTPKEDSQLVELITALGGCCRF LRDGYKTPKEDASLAVCTACAG PRAPKTAGGTTTPTSTSSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL LYDIRKTITETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE MILTITLETTPQFIDQQGFYLALDNKMIVE SKEPSINLLDSPHFRQENQYPSVRVEHLPRDQ SGCVDFFKALVQLUKFALVQGANVQTVSKVEHLPRDQ SGCVDFFKALVQLUKFALVQCHVMIVTLANDLENGSG SGVDFFKALVQLUKFALVQCANTATQLIDEA SEGDMSISILTCEMVYLAMYMRTQPGLAAC MILTITLETTPGFIDQGKTSACAGAG MILTITLETTPGFIDQGKTSACAGAG MILTITLETTPGFIDQGKTSACAGAG MILTITLETTPGFIDQGKTSACAGAG MILTITLETTPGFIDQGKTSACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA		1	1	1	ļ		
WHITASPPVHVASPPVASPP	1		1	1			
ASPPVHVASPPVJSGGDSTSDCFFP QPGAVFHSLAPSLGGWSHLVAALP  434 1784 A 3516 142 590 GGVNRPRSETEQVKTPVLISSWDYRHPPRPA SFFVFLV*TGFTALARMVLISWPCDLFTSASQ SAGTIGVRHARALLYFEQESHSVTQAGWVQ WHALGSLQPLSLEDRLSPGVLGCSALCRSGV RTKPGINMVTSRERGTTRLFKEG  435 1785 A 3529 1 3161 MSLVRAALEADELDLGVKGGPQSVHVLA DEVQRCQSILNSLLPRASTSKEVDASLLSVVS PAFAVEDSQLVELITKQGHTKLQGRYGCCRF LRDGYKTPKEDPINLLYVRPPABLKLFENIEC EWPLFWTYPLIDGVPSADSQLVGVKALEA VLIKGKNGVPLLPELYSVPDRVDEEYQNPHT VDRVPMGKLPHMWQQSLYLGSLMAEGFLA PGEIDELNRFSTVPKPAALDVGVPARLSHTYSSLEIF LPFLNSVSGCNRMKLSGRPYRHMGVLGTSK LYDIRKTITFTFTQFIDQQFYLALDNKMIVE MLRTDLSYLCSRWRMYGQFTTIFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDGFGKLYSEDYDDN YDYLESGNWMNDVDSTSHARCGDEVARVL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLFTKLFQA SRPSFNLLDSPHPRQENQYPSVRVEHLPRDQ SGEVDFKALVLQLKETSSLQEQADLLYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG ERRIWGLIRYISGLIRKSLARCSAEBATEGL MPLRTGLIQVMATELSILGSKEFGVERSVRPTD SNVSPALSHLIGNENGTSTRETRGTFGGLKSEIN QSPGTSMTPSSGSFPSAYDQQSSKDSRQGW QRRRILDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHWESVL NRVPQPEYRGLLVEAILVLALDHLADIENHSIGS LAVEKVVHANDLLFGQKTLGADDTMLAKD PASGICTLLVDSAPSGRFGTMTYLSKAAATY VQEELPHSLGAMQ VGERRHJCLARLLVLALDELHISIGS LAVEKVVHANDLLFGQKTLGADDTMLAKD PASGICTLLVDSAPSGRFGTMTYLSKAAATY VQEELPHSLGAMQ VGERRHJCLARLLVLALALDHLHILSCAEFF LATERTORM LASEL LATERTYMPOTORY		1		l	1		
OPGAVPPHSLAPSLGGWSHLVAALP		i		1			1
434		1	1	1			
SFFVELV*TOFTÄLARMVLISWPCDLPTSASQ SAGITGVRHHARLLYFEQESHSVTQAGWVVQ WHNLGSLQPLSLEDRLSPQVLGCSALCRSGV RTKFGRMVTSRERGTTRLPKEG  435 1785 A 3529 I 3161 MSLVRAALEALDELDLFGVKGGPQSVIHVLA DEVQHQQSLINSLLJPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEITIKLQGRYGCCKF LRDGYTPKEDPRNLYYENPAELKLFENEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHIMWGQSLYLLGSLMAEGFLA PGEIDPLNRRSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTISIRTTTRKITAPVSILAETEEIKTIL KDKGTYVETLAEVYPRVQARILSHYSSLEIF LPFLNSVSGCNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQFYVALDNKMIVE MIRTDLSYLCSKWRMTGGPTITTPISHSMLDE DGTSLNSSILAALRKMQDGVFYGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKL VSEDVDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLTIKLFQA SRPSSRILLDSHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNBERSATVRELLTELYGKVG EIRHWGLIRYISGILKKVEALDEACTDLLSH QKHLTVGLPPPEPEKTISAPLPYEALTQLIDEA BEGDMSISULTQEMYVLAMYMRTQPGLFAE MFLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHISGSKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTSSGSSPSAYDQSSKDSRQGQW QRRRRLDGALMRNPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTFGEIKFSVHVESVL NRVPQPEYRQLLVEAILVUTMLADIENHSIGS ILAVEKVHANDLFLQEGKTLGADDTMLAKD PASGICTILLYDSAPSGRGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CPPLTWELLEVKKAEVLQDSLDGRYSTPSSCL	İ	1 ,	1	1 -			
SAGITGVRIHIARILLYFEQESHSVTQAGWWQ WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV RTKFGRIMVTSRERGTTRLFKEG  435 1785 A 3529 1 3161 MSLVRAALEALDELDLFGVKGGFQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTRQEIITKLQGRYGCCRF LRDGYKTPKEDPNRLYJFRNPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWQGSJVLIGSLMAEGFLA PGEIDPLINRETSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTISIRTTRKITAPVSILAETEEKTIL KDKGYVETIAEVYPRVQPARLASHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTQFIDQQFYLALDNKMIVE MLRTDLSYLCSSWRMTGQPTITPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLITISCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHFKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SGRVDFKALVLQLKETISSLQEQADILYMLYT MKGPDWTELJYNERSATVRELITELYGKVG EIRHWGLIRYISGILRKEVEALDEACTDLLSH QKHLTVGLPPEPREKTISALFYPEALTDLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNNSPAISHEIGAVGATKTERTGIMQLKSEIK QSFGTSMTFSSGSFFSATVQQSSKDRQQW QRRRLDGALNRVPVGFYQKVMKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEFYRQLLVEALIVLTMLADIEINSIGS LIAVEKIVHIANDLFILQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ 436 1786 A 3546 73 393 CPPLTWELLEVKKAEVLQDSLDGRYSTPSSCL	434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA
WHNLGSLQPLSLEDRLSPGUGCSALCRSGV RTKFGINMVTSRERGTTRLPKEG  A 3529 1 3161 MSLVRAALEALDELDLFGVKGGFQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF LRDGYKTPKEDPNRLYYENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLIPELYSVPPDRVDEEYQPPHT VDRVPMGKLPHNWQGSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVQVYPSLPHGCS SKSPSHQCTIBIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPPLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQOFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTTTFPISHMLDE DGTSLNSSILAALRKMQDG YFGGARVQTGKL SEFLTTSCCTHLSFMDPGEFGKLYSEDYDDN YDYLESGNWMNDVDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPROENQVPSVRVEHILIRDQ SGEVDFKALVLQLKETSSLQFQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKVALDEACTDLSH QKHLTVGLPPPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEBATEGL MNLSPSAMKNLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFSTVROMYDQSSKDSRQQW QRRRLDGALNRVPVGFYQKVMKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEFYRLLVEALLVLTMLADIENHSIGS IIAVEKIYHANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CPPLTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	} ```	1	ł	i	ľ	SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ
WHNLGSLQPLSLEDRLSPGUGCSALCRSGV RTKFGINMVTSRERGTTRLPKEG  A 3529 1 3161 MSLVRAALEALDELDLFGVKGGFQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF LRDGYKTPKEDPNRLYYENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLIPELYSVPPDRVDEEYQPPHT VDRVPMGKLPHNWQGSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVQVYPSLPHGCS SKSPSHQCTIBIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPPLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQOFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTTTFPISHMLDE DGTSLNSSILAALRKMQDG YFGGARVQTGKL SEFLTTSCCTHLSFMDPGEFGKLYSEDYDDN YDYLESGNWMNDVDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPROENQVPSVRVEHILIRDQ SGEVDFKALVLQLKETSSLQFQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKVALDEACTDLSH QKHLTVGLPPPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEBATEGL MNLSPSAMKNLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFSTVROMYDQSSKDSRQQW QRRRLDGALNRVPVGFYQKVMKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEFYRLLVEALLVLTMLADIENHSIGS IIAVEKIYHANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CPPLTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1		]	ł	1	SAGITGVRHHA\RLLYFEOESHSVTQAGW\VQ
RTKFGINMVTSRERGTTRLPKEG  435 1785 A 3529 1 3161 MSLVRAALEALDELDLFGVKGGPQSVIHVLA DEVQHCQSLINSLIPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEITKLQGRYGCCRF LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA POEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTISRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTTFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDVDSTSHARCGDEVARYL DHLLAHTAPIFKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEHLIPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLRYISGILRKEVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISLITQEIMVYLAMYMRTQPGLFAE MFRLIGLIQVMATELAHSLRCSAEEATEGIL MNLSPSAMKNLHHILSGKEFOVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQQOW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTHGEIKFSVHVESVL NRVPQPEYRGLLVEAILVLTMLADIENISIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ 436 1786 A 3546 73 393 CPPLTWBLLEVKAEVLQDSLDGRYSTPSSCL		1		1	1		
435 1785 A 3529 I 3161 MSLVRAALEALDELDLFOVKGOPQSVHTVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVV FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNOVPLLPELYSVPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTISISTITRKITAPVSILAETEEIKTIL KDKGGTVETIAEVYPIRVQPARILSHYSSLEIF LPPLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISSISMLDE DGTSLNSSILAALRKMQDG YFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDDYDSTSHARCGDEVARYL DHLLAHTAPHVLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHRRQENQYPSVRVEHLPRDQ GGEVDFKALVQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGIRKKVEALDEACTDLLSH QKHLTVGLPPEREKTISAPLPYEALTQLIDEA SEGDMSISILTQEMVYLAMYMRTOPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MMLSPSAMKNILHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELVKAEVLQDSLDGRYSTPSSCL	1		1	ł			
DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEITKLQGRYGCCRF LRDGYKTPKEDPNRLYYENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKTAPVSILAETEEIKTL KDKGFLYETTIAEVYPIRVQPARILSHTYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTITTFTPGPIDQQFYYALDNKMIVE MLRTDLSYLCSRWRMTGQPTITTFISHSIMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLITISCCTHLISFRDPOPEGKLYSEDYDDN YDYLESGNWMDYDSTSHARGGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHQENGVPSVRVEHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELITELYGKVG EIRHWGLRYVLAMYMRTQPGLFAE MFRLRIGLIQVMATELAFISLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSGSFPSAYDQSSKDSRQGW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTFGEIKFSVHVESVL NRVPQFEYRQLLVEALILVTMLADIEINSIGS IIAVEKIVHIANDLFILQEQKTLGADDTMLAKD PASGICTILYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELVKAEVLQDSLDGRYSTPSSCL	40.5	1505	<del>  </del>	2620	<del> </del>	2161	
FPAFAVEDSQLVELTKQEITIKLQGRYGCCRF LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHNWGQSLYILGSLMAEGFLA PGEIDPLNRSTYVRFDVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGTYVETTAEVYPRVQFARILSHTYSSLEIF LPPLNSVSGENRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFQFIDQQGFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETISSLQEADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MMLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTITREMTPGEIKFSVHVESVL NRVPQPEPYGLLVEALLVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTILLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	435	17/85	A	3329	1	3101	
LRDGYKTPKEDPNRLYV/ENPAELKL/FENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVFKPDVVVQYPSLPHGCS SKSPSHQCTIISRTTRKITAVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRKLSGRPYRHMGVLGTSK LYDRKTIFTFTQFIDQOQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSLAALRKMQDGYFGGARVQTGKL SEFLTISCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHILAHTAPHFKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLDSPHPRQENQVPSVRVEHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSSAISHEIGAVQATKTERTGIMQLKSEIK QSPGTSMTSSGSFFSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEPYROLLVEAILVITMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTILLYDSAPSGRFGTMTYLSKAAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1		1	1			
EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLIPLYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTTIFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMMDYDSTSHARCGDEVARYL DHILLAHAPHPKLAPTSQKGGLDRPQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEMVYLAMYMRTQPGLFAE MFRLIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNILHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL,VLTMLADIENISIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL			1	1		1	
VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMVGQSLYILGSLMAEGFLA PGEIDPLNRFSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVOPARILSHYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTTPOFIDQQOFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQYPSVRVEHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELTNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPSTRQLLVEAILVLTMLADIEINISIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEELPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1	l	1		}	
VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGGIIDPLNRRTSTVPKPDVVQVYPSLPHGCS SKSPSHQCTIISRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVGCNNRMLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVMKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTILLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CPPLTWELLEVKKAEVLQDSLDGRYSTPSSCL	1			1	İ		7 7
PGEIDPLNRRFSTVPKPDVVQVYPSLPHGCS SKSPSHQCTIISRTTTRKITAPVSILAETEELKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFPQFIDQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTIFFISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEHLLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MMLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CPPLTWELLEVKKAEVLQDSLDGRYSTPSSCL		1	1	1	i	1	VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT
SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPPINSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTTPQFIDQQGYLALDNKMIVE MLRTDLSYLCSR WRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLITSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGIRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL		1		1	1		VDRVPMGKLPHMWGQSLYILGSLMAEGFLA
SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPPINSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTTPQFIDQQGYLALDNKMIVE MLRTDLSYLCSR WRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLITSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGIRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL		ļ	1	1	1	1	PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS
KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFPGIDQQGFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTTIFFISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCTHLSFMDPGPEGKL YSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	}	1				1	SKSPSHOCTIISIRTTRKITAPVSILAETEEIKTIL
LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPOFIDQQOFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHILLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEHHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMYVLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL		[	1				
LYDIRKTIFTFIPQFIDQQQFYLALDNKMIVE MLRTDLSYLCSR WRMTGQPTITFISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEIHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL			i .	1 .	1		
MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGFEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL				1	1		
DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIE\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL			1	1	l		
SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATTY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL			1		1 -		
YDYLESGNWMNDYDSTSHARCGDEVARYL DHILAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINSIGS IIAVEKIVHIANDLFILQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1		·			
DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADLIYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL		1	1	1		1	
TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	· ·	ł	}			1	
SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEINISIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL						1	
SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEINISIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL			İ				TCDLMSLVTKAKELHVQNVHMYLPTKLFQA
SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MMLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1		1			
MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEINHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1					
EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL			1	1	.]		
QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1	1				
SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1		1				
MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEINSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1	1	1		1	
MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFILQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL				Ī	ľ		SEGUMSISILI QEIMV YLAMYMKI QPGLFAE
SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTITREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1	1	1			
QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1			1	1	1	
QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1		.		l	
QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1		1	1	1		QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW
GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1			1	1	1	
NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL				1	1.	1	
IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1		1	1	1 .		
PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ  436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL	1	1	1	1	1	1	
VQEFLPHSICAMQ	1			l	1	1	
436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL		1		1	1	1	
						<u> </u>	
	436	1786	A	3546	73	393	
1 1 1 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	1				l		EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KGKGKTIRGI*TFKGRKGGTYQREHDANPLA PXSARSCWMRKG
437	1787	A	3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG SASGCRYAAHPHWGLGGAAAAGGSWEPQPP RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA AQPGARARTSPPPASARNMAARPAATLAWSL LLLSSALLREGCRARFVAERDSEDDGEEPVVF PESPLQSPTVLVAVLARNAAHTLPHFLGCLER LDYPKSRMAIWAATDHNVDNTTEIFREWLK NVQRLYHYVEWRPMDEPESYPDEIGPKHWP TSRFAHVMKLRQAALRTAREKWSDYILFIDV DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS NFWCGITPKGFYKRTPDYVQIREWKRTGCFP VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK LMDNIDQAQLDWELIYIGRKRMQVKEPEKA VPNVANLVEADYSYWTLGYVISLEGAQKLV GANPFGKMLPVDEFLPVMYNKHPVAEYKEY YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE TSTIWDNETVATDWDRTHAWKSRKQSRIYSN AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT TLG
439	1789	A	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLSTLTGRSY GQPSLQDELKDNTTVFTRILDRLLDGYDNRL RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI DVFFRQSWKDERLKFKSPMTVLRLNNLMAS KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE DGTLLYTMRLTVR\AECPMAFGRDFFMD\AH ACPLKFGSYAYTRAEVVYEWTREPARSVVV AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV MTIHFHLKRKIGYFVIQTYLPCIMTVILSQVSF WLNRESVPARTVFGVTTVLTMTTLSISARNSL PKVAYATAMDWFIAVCYAFVFSALIEFATVN YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN NTYAPTATSYTPNLARGDPGLATIAKSATIEP KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL FGIFNLVYWATYLNREPQLKAPTPHQ STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK
440	1790	A	3568	1	350	FWEVISDEHGIDPTGTYHGDSDLQLERINVYY NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569	2	1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ FIEELITKWQKNDQELISDPLQQCFKKDEILDG QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE VKTPFDLAKAQENSNSVKKKTKFVNLYTREG QDRLAVLLPGRHPCDCLGQKHKLINNCLICG RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS NKSQKLLKKLMSGVENSGKVDISTKDLLPH QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

		<b>.</b>	650	<b>5</b> 10	Design of the second	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino	
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
I				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			·	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						DDESDYFASDSNQWLSKLERETLQKREEELR
				1		ELRHASRLSKKVTIDFAGRKILEEENSLAEYH
1			l			SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL
					Į.	VNPNMYQSPPQWVDHTGAASQKKAFRSSGF
			1	i	i	GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH
			•	1		QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
	i	l	ļ			TAKKPSPQEVSELQATYRLLRGKDVEFPNDY
		İ	1			PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS
	ļ	i		l	ļ	PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH
					•	OGAKKGLMKQNKAV
442	1792	A	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK
442	1772	^	3370	1 *	-01/	KTAGVKPYECTICGKAFMRLSSLTRHMRSHT
	İ			1		AIRANEKPYKCKEC\GRAFSLSQILSK\HERSH
		•				TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK
	1	Į.			1	PYECKQCGKALSCSSSLRVHERIHTGEKPYEC
	1	l	1		1.	KQCGKAFSCSSSIRVHERTHTGEKPYACK\EC
		1				GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA
			1	Į.	1	FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS
	1	1	1			TOTOTTE DITTOEVE VICE CANCELLAND LEDV
	1	İ	1			TSIQIHERIHTGEKPYKCKECGKSFSARPÄFRV
<u> </u>	ļ			1	1	HVRVHTGEKPYKCKECGKAFSRISYFRIHERT
!	1	1				HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG
	1				_	EKPYECKECAKTFISLENFRRHMITHTGDGPY
1	1	ļ .	1		1	KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ
İ	ì	ŀ				CGKAFSCSSYIRIHKRTHTGEK\PYECKECGK
	l		1		Į.	AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS
						LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS
İ		1	1			LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF
	1	1		1		QSCENSH*REKSCQCK*YRKRDTR*FMYSQV
						PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL
ľ			1	-		VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
		ļ				DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL
1		1				ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK
443	1795	ļ ^	33,0	20,	1	LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT
444	1/34	A	3302	3333	1,707	MTKVTLENFYSNLIAQHEEREMRQKKLEKV
1		1	I		1	MEEEGLKDEEKRLRRSAHARKETEFLRLKRT
		ł	1			RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH
1		1				VYAMKILRKADMLEKEQVGHIRAERDILVEA
	1			1	1	DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM
		1		1		MTLLMKKDTLTEETQFYIAETVLAIDSIHQL
	1	1		1		GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK
	1	1			1	
	1	1	1	1	1	KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE
}	1		1	1	1	TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN
	1.	1	1	1	1	KLCDWWSLGVIMYEMLIGYPPFCSETPQETY
	1	1	1		1	KKVMNWKETLTFPPEVPISEKAKDLILRFCCE
	1	1	1	1	1	WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA
		1	1			AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE
	1				1	TDYKNKDWVFINYTYKRFEGLTARGAIPSYM
		1			1	KAAK
445	1795	A	3584	1	6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD
1 773	1 ***	1.,		Ī -	1	GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK
1	1		1			NHRLYFLSAASRPLCSGGHASNKEKEMVTSL
		1				
		1				FCKLGVLVRHRISLFGNDATSIVNCLHILGOT
						FCKLGVLVRHRISLFGNDATSIVNCLHILGQT
						LDARTVMKTGLESVKSALRAFLDNAAEDLE
						LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKOGOFTHTRNOPKGVTQIINYTTVA
						LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENILKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGOHOFGEDLILEDVQVSCYRI
						LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA
						LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENILKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGOHOFGEDLILEDVQVSCYRI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		MPHVMEVILPMLCSYMSRWWEHGPENNPER AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM EKLKKKAATVVSBEDHLKAEARGDMSEAEL LILDEFTTLARDLYAFYPLLIRPGDYNRAKWL
						KEPNPEAEELFRMVAEVFIYWSKSHNFKREE QNFVVQNEINNMSFLITDTKSKMSKAAVSDQ ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
			-		·	DPEKTVERVLDIANVLFHLEQKSKRVGRRHY CLVEHPQRSKKAVWHKLLSKQRKRAVVACF RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ LILLFSRTALTEKCKLEEDFLYMAYADIMAKS
						CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM
	-		-			VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ
						KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP
						FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI
						FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
						KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK
						LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH
						NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
446	1796	A	3592	1	355	FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
						DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML PRLVSNSWTQAILLPRPPKMLGLQV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LFVGGGPICPEGASGFAPGPAPARVGVDAEV GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI
						NGTLALGLKP**AWGWGEWRPKG
448	1798	A	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM RITNENFVDAYENSNSTEFVSLASKVKDALKL LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE FSIPQHLVEEAERVMAEERVVMLPPRARSLKS FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR GVELMRFTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVTVYNTL SPMEPHALVQLCGTYPPSYNLTFHSSQNVL LITLITNTERRHPGIFEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHVKVRFKFFYLLEPGVPAGTCPKDYVEING EKYCGERSQFVVTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPVAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY
449	1799	A	3618	2	613	TRLPLFRDWIKENTGV FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE
450	1800	A	3620		2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE HRETYQKLLEDIAVLHRLAARLSSRAEVVGA

NO: of nucleotide peptide cotide sequence uence used and provided in peptide sequence uence used and peptide sequence used	TARS PGQ ALLE TKA EEEQ
eotide sequence    Sequence   Seq	TARS PGQ ALLE TKA EEEQ
	TARS PGQ ALLE TKA EEEQ
914   ng to first amino acid residue of peptide residue of peptide sequence   1 Threonine, V=Valine, W=Tryptopham, Y=Tyrosine, X=Unknown, *=Stop codon,  -possible nucleotide deletion, \ -possible nucleotide insertion   VRQEKRMSKATEVMMQYVENLKRTYF   AELMEFKKLANQNSSRSCOPSEDGVLR   MSLTLGKNMPRRFVSVAVVPKFNALNI TPSSSSIPSI, PAL SESPNGKGSLPVTSALP   NGKTNGDPDCEASAPALTLSCLEELSQI RMEERAYSKGFOGGILKYTKELQDLKEE   KSESPEPPEEVETEETEKDPSSSKLEEL   QVMYPYLCQHWQVIWMMAAVMLVLT   GLYNSYNSCAEQADGPLGRSTCSAAQK   WSSGLQHEQPTEQ   QVMYPYLQCHWQVIWMMAAVMLVLT   GLYNSYNSCAEQADGPLGRSTCSAAQK   WSSGLQHEQPTEQ   QVMYPYLQCHWQVIWMMAAVMLVLT   GLYNSYNSCAEQADGPLGRSTCSAAQK   WSSGLQHEQPTEQ   QVMYPYLQCHEALKGRGAFNSQL   RIHTOERPYQCKELKGRGAFNSQL   RIHTOERPYQCKELKGRGAFNSQL   RIHTOERPYQCKELKGRGAFNSQL   RIHTOERPYQCKELKGRGAFNSQL   RIHTOERPYQCKELKGRGAFNSQL   QCTALGVWTAPAPVCIAVQCQHLEALMG*PVPTAFAYGSSCKYECHTVYRVI   MLHSRGCYLWNGHFTT*EALSCEPLER   V*CSS*CEEGFALIGFEVVQCTALGVW   V*CIAVQCQHLEALNEGTMG   QCTALGVWTAPAPVCIAVQCQHLEALMG*GFGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	TARS PGQ ALLE TKA EEEQ
amino acid residue of peptide sequence	TARS PGQ ALLE TKA EEEQ
residue of peptide sequence	TARS PGQ ALLE TKA EEEQ
peptide sequence	TARS PGQ ALLE TKA EEEQ
Sequence	TARS PGQ ALLE TKA EEEQ
VRQEKRMSKATEVMMQYVENILKRTYF	TARS PGQ ALLE TKA EEEQ
MSI.TLGKIMPRRRVSVAVVPKINALNI	PGQ ALLE TKA EEEQ
TPSSSIPSLPALSESPNGKGSLPVTSALE   NGKTNGDPDCEASAPALTILSCLELSQ  RMEEAYSKGPQEGLKKTKELQDLKEE   KSESPEEPEVEETEEEKDPRSSKLEEL   QVMYPKL.QHWQVIVMMAAVML.VL]   GLYNSYNSCAEQADGPLGRSTCSAAQK   WSGQHEQPTEQ	ALLE TKA EEEQ
NGKTNGDPDCBASAPALTLSCLEELSQE   RMEBEAYSKGFQEGLKKTKELQDLKEE   KSESFEPEPEVETEEEEEKDPRSSKLEEL   QVMYPKLCQHWQVIWMMAAVMLVLT   GLYNSYNSCABQADGPLGRSTCSAAQK   WSGLQHEQPTEQ	TKA EEEQ
RMEEAYSKGPQEJLKKTKELQDLKEE   KSESPEPPEVEBTEEEKDPRSSKLEEL   QVMYPKLCQHWQVIWMMAAVMLVLT   GLYNSYNSCAEQADGPLGRSTCSAAQK   WSSGLQHEQPTEQ	EEEQ
KSESPEPEVETTEEEKOPRSSKLEEL	
QVMYPKLCQHWQVIWMMAAVMLVLTGLYNSYNSCAEQADGPLGRSTCSAAQK	VDD
1801   A   3623   504   198   QLIQHOTVHTGKLYECKECGKAFNOO RHQRIHTGEKPYECKVCGKAFRVSQL RIHTGERPYCKVCGKAFRVSQL RIHTGERPYCKECKCGKAFRVSQL RIHTGERPYCKECKCGKAFRVSQL RIHTGERPYCKECKCGKAFRVSQL RIHTGERPYCKECKCGKAFRVSQL RIHTGERPYCKELKGGAEMLAVLAY NRTPVNYGK   452   1802   A   3628   2   195   MTCLHSAKAFHY*SSCSFSCEEGFALIG QCTALGVWTAPAPVCIAVQCQHLEALM MG*DYPFTAFAYGSSCKYECHTVYRVV MLHSRGCYLWNGHFTT*EAISCEPLERI V*CSFSCEEGFALIGPEVVQCTALGVW VCIAVQCQHLEALNEGTMG   V*CSFSCEEGFALIGPEVVQCTALGVW VCIAVQCQHLEALNEGTMG   10AKGLGIWHVPNKSPMQHWRIKGSLL DTGFLQTLGHNLLGIYQKYPVKYGEGF AGFVQFRVFNNERAANLCAGMRVTC HHCIGGGGYFEAAVLLFYR   1804   A   3641   1   362   TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPRVW ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLLINCDI GCLYHDVMLETLTLISSLGKVLLINCDI GCLYHDVMLETLTLISSLGKVLLINCDI SLLLGFTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRNDTT SRVRAPSYDDIT   SRVRAPSYDDIT   SRVRAPSYDDIT   SRVRAPSYDDIT   SRVRAPSYDDIT   SRVRAPSYDDIT   LLELKNIHNLG*AKFFLN*IQKALKKRCTIVWEQ	
WSSGLQHEQPTEQ	AAL
1801	D2M
RHQRIHTGEKPYECKVCGKAFRVSSQL   RIHTGERPYQCKELKGRGAEMLAVLAV   NRTPVNYGK     A   3628   2   195   MTCLHSAKAFHY*SSCSFSCEEGFALIG   QCTALGVWTAPAPVCIAVQCQHLEALM   MG*DYPFTAFAYGSSCKYECHTVYRV   MLHSRGCYLWNGHFTT*EAISCEPLERI   V*CSFSCEEGFALIGPEVVQCTALGVW   VCIAVQCQHLEALNEGTMG     453   1803   A   3637   662   142   IQAKGLGIWHVPNKSPMQHWRIKGSLL   DTGFLQTLGHNLLGIYQKYPVKYGEGF   DNGPVIPVVYDFGDAQKTASYYSPYGC   AGFVQFRVFNNERAANALCAGMRVTC   HHCIGGGGYFPEASPQQCGDFSGFDWS   HVGYSSSREITEAAVLLFYR     454   1804   A   3641   1   362   TQVHPAMLGLDELGRSGCGHCTQADL   AAGRDPGQDNDRNTAEPAFPPPPRVM   ALRAPAQSSVTFEDVAVNFSLEEWSLL   GCLYHDVMLETLTLISSLGKVLILNCDI   SLLLGPTFNSCQVSSQPPRVAGLGLPLK   RPQPPSPRGFRTVRAGVPGAHPQDTPC   PRKVPLVGEAPGLPPEERSRGWRRDTF   SRVRAPSYDDIT     456   1806   A   3656   396   8   QIVSFNSYLTLYTKNNLKSMKDLNVM   SRVRAPSYDDIT   SRVRAPSYDI	COTT T
RIHTGERPYQCKELKGRGAEMLAVLAY   NRTPVNYGK	SIL
NRTPVNYGK	VEA
1802   A   3628   2   195   MTCLHSAKAFHY*SSCSFSCEEGFALIG QCTALGVWTAPAPVCIAVQCQHLEALM MG*DYPFTAFAYGSSCKYECHTYTRVI MLHRRGCYLWNGHFTT*EAISCEPLERI V*CSFSCEEGFALIGPEVVQCTALGVW VCIAVQCQHLEALNEGTMG	Vav
QCTALGVWTAPAPVCIAVQCQHLEALM   MG*DYPFTAFAYGSSCKYECHTVYRVE   MI.HSRGCYLWNGHFTT*EAISCEPLER   V*CSFSCEEGFALIGPEVVQCTALGVW   VCIAVQCQHLEALNEGTMG   IQAKGLGIWHVPNKSPMQHWRKGSLL   DTGFLQTLGHNLLGIYQKYPVKYGEGF   DNGPVIPVVYDFGDAQKTASYYSPYGG   AGFVQFRVFNNERAANALCAGMRVTG   HHCIGGGGYFPEASPQQCGDFSGFDWS   HVGYSSSREITEAAVLLFYR     454	5EVV
MG*DYPFTAFAYGSSCKYECHTVYRVI   MLHSRGCYLWNGHFTT*EAISCEPLER	
MLHSRGCYLWNGHFTT*EAISCEPLERI	CLD
V*CSFSCEEGFALIGPEVVQCTALGVW. VCIAVQCQHLEALNEGTMG  453 1803 A 3637 662 142 1QAKGLGIWHVPNKSPMQHWRIKGSLI. DTGFLQTLGHNLLGIYQKYPVKYGEGF DNGPVIPVVDFDAQKKTASYYSPYGG AGFVQFRVFNNERAANALCAGMRVTG HHCIGGGGYFPEASPQQCGDFSGFDWS HVGYSSSREITEAAVLLFYR  454 1804 A 3641 1 362 TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPPRVM. ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI AAAGRGASGALTGEGGGEQGRRVGLG SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVM LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
453 1803 A 3637 662 142 IQAKGLGIWHVPNKSPMQHWRIKGSLL DTGFLQTLGHNLLGIYQKYPVKYGEGK DNGPVIPVYDFGDAQKTASYYSPYGG AGFVQFRVFNNERAANALCAGMRVTG HHCIGGGGYFPEASPQQCGDFSGFDWS HVGYSSSREITEAAVLLFYR  454 1804 A 3641 1 362 TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPPRVM. ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI AAAGRGASGALTGEGGGEQGRRVGLG SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVM LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
453 1803 A 3637 662 142 IQAKGLGIWHVPNKSPMQHWRIKGSLI DTGFLQTLGHNLLGIYQKYPVKYGEGR DNGPVIPVYDFGDAQKTASYYSPYGG AGFVQFRVFNNERAANALCAGMRVTG HHCIGGGGYFPEASPQQCGDFSGFDWS HVVGYSSSREITEAAVLLFYR  454 1804 A 3641 1 362 TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPPRVM ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI AAAGRGASGALTGEGGGEQGRRVGLG SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
DTGFLQTLGHNLLGIYQKYPVKYGEGE DNGPVIPVYDFGDAQKTASYYSPYGG AGFVQFRVFNNERAANALCAGMRVTG HHCIGGGGYFPEASPQCGDFSGFDWS WVGYSSSREITEAAVLLFYR TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPRVM ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	RYRT
DNGPVIPVVYDFGDAQKTASYYSPYGG AGFVQFRVFNNERAANALCAGMRVTC HHCIGGGGYFEASPQQCGDFSGFDWS WVGYSSSREITEAAVLLFYR  454  1804  A 3641  1 362  TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPPRVM ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456  1806  A 3656  396  8  QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
AGFVQFRVFNNERAANALCAGMRVTC HHCIGGGGYFPEASPQQCGDFSGFDWS UHVGYSSSREITEAAVLLFYR  454  1804  A 3641  1 362  TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPRVM. ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRRDTP SRVRAPSYDDIT  456  1806  A 3656  396  8 QIVSFNSYLTLYTKNNLKSMKDLNVM LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
HHCIGGGGYFPEASPQQCGDFSGFDWS WYGYSSSREITE\AAVLLFYR  454 1804 A 3641 1 362 TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPRVM. ALRAPAQSSVIFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI SLLLGPTFNSCQVSSQPPRVGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
454 1804 A 3641 1 362 TQVHPAMLGLDELGRSGCGHCTQADL AAGRDPGQDNDRNTAEPAFPPPPRVM ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI 455 1805 A 3646 2 414 AAAGRGASGALTGEGGGEQGRRVGLC SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRRDTP SRVRAPSYDDIT 456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
AAGRDPGQDNDRNTAEPAFPPPPRVM. ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI AAAGRGASGALTGEGGGEQGRRVGLC SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
AAGRDPGQDNDRNTAEPAFPPPPRVM. ALRAPAQSSVTFEDVAVNFSLEEWSLL GCLYHDVMLETLTLISSLGKVLILNCDI GCLYHDVMLETLTLISSLGKVLILNCDI AAAGRGASGALTGEGGGEQGRRVGLC SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	RFGD
455 1805 A 3646 2 414 AAAGRGASGALTGEGGGEQGRRVGLC SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRRDTP SRVRAPSYDDIT 456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	<b>LAAA</b>
455 1805 A 3646 2 414 AAAGRGASGALTGEGGGEQGRRVGLO SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	NEAQ
SLLLGPTFNSCQVSSQPPRVAGLGLPLK RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
RPQPPSPRGPRTVRAGVPGAHPQDTPC PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
PRKVPLVGEAPGLPPEERSRGWRRDTP SRVRAPSYDDIT  456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVNT LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
456 1806 A 3656 396 8 QIVSFNSYLTLYTKNNLKSMKDLNVN LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	GLQE
LLELKNIHNLG*AKFFLN*IQKALIKRK P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
P/LIKIK/SFCSLSDTIKKMKRQTIVWEQ	
SVKELVSKI <u>Y</u> EAFLQFNKTVNRPVFDIR	
	veáv
457 1807 A 3660 14 1961 SEAKLGGPTGMDLWQLLLTLALAGSS	)ARCC
457 1807 A 3660 14 1961 SEAKLGGPTGMDLWQLLLTLALAGSS SEATAAILSRAPWSLQSVNPGLKTNSS	
TKCRSPERETFSCHWTDEVHHGTKNLO	
YTRRNTQEWTQEWKECPDYVSAGEN	
SSFTSIWIPYCIKLTSNGGTVDEKCFSV	EIVO
PDPPIALNWTLLNVSLTGIHADIQVRW	
ADIOKGWMVLEYELQYKEVNETKWK	
ILTTSVPVYSLKVDKEYEVRVRSKQRA	
GEFSEVLYVTLPQMSQFTCEEDFYFPW	LLIПF
GIFGLTVMLFVFLFSKQQRIKMLILPPV	PVPKI
KGIDPDLLKEGKLEEVNTILAHDSYKI	
DDSWVEFIELDIDEPDEKTEESDTDRL	
EKLHINLGVKDGDSGRTSCCEPDILET	EFHS
DIHEGTSEVAQPQRLKGEADLLCLDQI	EFHS SSDH
SPYHDACPATQOPSVIQAEKNKPQPLF	EFHS SSDH DFNAH
STHQAAHIQLSNPSSLSNIDFYAQVSDI	EFHS SSDH OFNAH NQNN TEGAL
VVLSPGQKNKAGMSQCDMHPEMVSL	EFHS SSDH OFNAH NQNN TEGAL
LMDNAYFCEADAKKCIPVAPHIKVES	EFHS SSDH OFNAH INQNN TEGAL TPAGS
LNQEDIYITTESLT\TAAGSMGTGEHVI	EFHS SSDH OFNAH INQNN TEGAI TPAGS CQENI

PCT/US01/03800

	Cro m	1 Xat	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid,
nucl-	peptide	1104	in in	nucleotide	Iocation	F=Phenylalanine, G=Glycine, H=Histidine,
	4	!	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	nence	l	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence	Į.	ł	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	j	}	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	<b>{</b>	l .			sequence	/=possible nucleotide deletion, \=possible
			1	peptide	ļ	nucleotide insertion
		ļ.,		sequence	<u> </u>	PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS
		1	]		ļ	SCGYVSTDQLNKIMP
	ļ				120	TRAPASGRSGAGLALSANAPDSGGHPGATEG
458	1808	Α	3663	154	462	PAGSLAHASGSARGTWRVRGRGSHGWERTV
	ì	1	1	ì	Ì	GAGGCANPVPALHSCASAPRGTGRVSALGPK
	}	)	1	}	1	
•	1	1			<u></u>	TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ
	1	Į.	1	1		SQNALGKYNTSMALFESNSFEKTILESPYYVD
	1	1	ļ		1	LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD
		1	1			FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF
		1	1			QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC
	1	i i	1			INQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR
	1	1	1	}		SAINGNSGFOHETHAEETPNQPFNSVHLFSFM
	1	-				VLALNVVTVATITVRHFVNQRADYQ\YQKLQ
	1	1	ì			NY
	1	+	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P
460	1810	Α	3670	650	1 337	K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA
1	[		1		1	GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI
ļ			l	1	1	S
l	<u>. }</u>			<del></del>	2000	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\
461	1811	A	3671	2472	2099	TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV
1	1	1		1		VQTGL*LLALSNPPALASQIAGISGMSHRAWP
1	İ	1	i			VOIGL*LLALSNPPALASQIAGISGISIAAAWI
	}	1	1			GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	A	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQT\N
'	1	1	,	1		CYYD/STKSFFYISCG*K\RKPTWAENRRLNA
1 .	- }	1		ı	1	KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG
ì	1	1	- 1	ļ	ł	HGS
463	1813	A	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR
403	1013	1 ^	1 50.5	1 3.3		KSSFSRDODVW*SOAVPKRQ*QRNPFSAGHP
	1	1	i	i		ORPPTSGSOSELLAQPRLRPGRKSSFSRDQDV
}	j .	j.	· I	1	1	WPGOKPRPSOOOHOMCASPTLGQRSPFALEP
1	ļ	l l	1			VPAYHGGRDPFASARPSPVGIPKPRAAPAGG
1	ł		Į.		1	GWRRIRPKSSTK
\		<del></del>	- 2000	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR
464	1814	A	3676	2253	320	KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM
1	1	1	1	<b>\</b>	1	RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS
1	ì	1	-			QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE
,	- 1	1	ł	1		CACHACONSINILA AND ACCHACION CONTINUE AND A VE
1	4		1	1	1	APACRISFLPLTRLRRTESVPSDINNPVDRAAE
1	1	1		1		PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT
l	-	1		1		FSTPSSPAPFPTSSNPSSATTPP\NPSP\GQR\DSR
1	1	1	1	1		FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE
1		i	- (	1	J	AADGTRLDDQPKADVLEAHEAEAEEPEAGK
			- 1	[	l	SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS
1		1	1	1	[	VYLOEWDIPFEQVELGEPIGQGRWGRVHRGR
· 1				1 .		WHGEVAIRLLEMDGHNODHLKLFKKEVMN
1	1	1			1	YROTRHENVVLFMGACMNPPHLAIITSFCKG
1		- 1	1	1	1	RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY
1		}		ł	1	LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF
1	1	-	1	1	1	\GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR
	1	- 1		ì	1	EMTPGKDEDQLPFSKAADVYAFGTVWYELQ
			1	1	1	ARDWPLKNQAAEASIWQIGSGEGMKRVLTS
j.		1		1	ł	VSLGKEVSENLSACWAFDLQERPS\FSLLMD
1		- 1	i	- 1		A STATE ASSESSED OF STREET ASSESSED ASSESSED OF STREET ASSESSED OF STR
1	1	1		1 .	Ì	MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR
	1					FERFGLGVLESSNPKM
465	1815	A	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY
1400	1013	^	1 30,7	1		WACVLOTHRAFCASNTEDLETVVNHIKHRYP
		1	ì	1	(	OAPLLAVGISFGGILVLNHLAQARQAAGLVA
		1	1	1	1	ALTLSACWDSFETTRSLETPLNSLLFNQPLTA
1		1	}	1		GLCQLVERLSY/E*DLQARTIRQFDERYTSVA
					1	ODOVD   DANDO   OD ON DESCRIPTION OF THE PROPERTY OF THE PRO

PCT/US01/03800

WO 01/57188

			1 000	Destinad	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning		F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	Ì	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	İ	09/496	correspondi	to last amino	
uence	i	1 .	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1	i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ì	1	1	peptide	1 -	/-possible nucleotide deletion, \-possible
	l .	(	1	sequence		nucleotide insertion
	<b></b>	<b>└</b> ──	<del></del>	Sequence		FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA
		1	1			DDPFSTVCALPKQAAQHSPYVALLITARGGHI
	1	Į.	1			GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
		1	1	ł	1	GFLEGLLPWQRW INSKLERQ I AKAIPQDFE
	1	ļ		ł		GLPDLRALLPSEDRNS
466	1816	A	3684	3	307	SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS
400	1010	1		1		ANISSQTGEARGQWPSVFKVLKEKKLSTKKS
	ļ	1	ļ		1	FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
	ì	i .	ì	1	1	TGVLOG
	ļ			0465	027	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA
467	1817	Α	3687	2465	837	RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ
	1	1		1	1	KLOOLEUME I VENEVA Y BEEEGGE TEBRIE BY OF
	1	1	1	1	1	DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF
	1		1		}	FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE
	1	1	Ţ	1		APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG
	1			1		MARARLAQLVRLAGGHCRRDTLWKRLFLLE
	i	1	ł	Į.	}	PPGPDRLRLGGRLALAELEELLEAVHAKSIGD
		1	1	1		IDPOLDCFLSMTVSWYQSLIKVLLSRFPQSCR
		-		1		HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH
}		1	1		1	LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY
	Ì	1	l	1	1	
	1		Ĭ			HHLESVINTACFTLWTRLL*GSGLDH*MSLFL
ł	i	1	1		1	ESWAYQIACQRQD*PALLGPRASQTLSDTKG
ļ .		١	1	1	1	FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR
	i	į	1		Ì	ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ
		1	1		1	TPPRAPLPESCPL\PLTTVSHLCPLSLRVFTSHL
1		1	1		ļ	DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA
	1	i	ì	1	.1	LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG
1	l l	1	1		1	
						GKQQRN
468	1818	A.	3691	960	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI
İ	1	1		Į.		QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH
	ı	1	İ	1	ł	DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF
1	1	)	Į.	1		PGNF\*FLVKTGFPHVGQTGFELLTSSDLAPLA
1	1	1	1	ı		SONGGITGMSPCAWPFFFFFFGLC
160	1010	TA-	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP
469	1819	^	3/14	7/7/	"""	HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF
		ì		i	í	VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL
1	İ	İ		1		ALIVATOR Y DATCESDENI DADONIANA LOCANO
	1	Ì	· .	J	]	SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG
1		1		ì	1	SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL
	1	-		Į.		ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK
		1	l		1	RSGHVNIVEPSLMLLKGSLQPGMWESTWQK
1		-		1	1	NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR
1		1	1		1	ERYHAADVNFNSGKIWSTTTAFPYOLFSKTK
1	1	1		[	(	FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH
	l	1	1 .	i		FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS
1				1	1	HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE
1			1	1	Į.	DINORIA MOLE PER GRUNNIODOGI ETI TONO
1	1 .	-	- 1	1	1	DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY
			1	1		DFHLKYLLKTQENVYNIIEEVKKICSVLGCVE
1		1	1			
			Ì		1	TKQITDAVNELSLILQRKGENFYQSSETSAKG
						LIEKVTTELSTSIYOLINVYCNSFYADFQPVNV
					1	LIEKVTTELSTSIYOLINVYCNSFYADFQPVNV
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL
-						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSOPSPVTLQIDFPATGWEYMKPD
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNI.EPPLKECIKHIARLSQKQTPLLLSE
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATTNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW
-						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSOPLEALGLLTSSFP
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSOPLEALGLLTSSFP
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DOEIRKVAVOOLDNLLNDELLEYLPQLVQAV
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVOLLLHRSLQSIQVAHRLYWL
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYOKLLAALQFCAGKALN
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEOKLIKILGDIGERVKSASDHQRQEVL
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEQKLIKILGDIGERVKSASDHQRQEVL KKEIGRLEEFFODVNTCHLPLNPALCIKGIDH
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEOKLIKILGDIGERVKSASDHQRQEVL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MIIYRCLSTGKDQRLVQMVPDAYTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EMEYFITEGGKNPQHFQDFVELCCRAYNIIR KHSQLLLINLLEMMLYAGLPELSGIQDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMLLSNPIW
470	1820	A	3718	430	75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVQLLLSLGSVFLWALL
471	1821	A	3723	891	494	LROSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	A	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIP LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824	A	3753		. 5262	RPLFAREGGIYAVLVCMQEYKTSV\LVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLLNLLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSP\ERAALETPIIQGQDGSPELLIRSLV GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG LELATTFEHFYQHYMADRILSFGSSWLEGAV LEQIGLCFPNRLPQLMLQSISTSEELQRQFHLP QLQRLDKLFLEQEDEEEKRL*EEEEEEEEEA EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG PHRRLQWTWLGRAELQFGKQILHVSTVQMW LLLKFNQTEEVSVETLLKDSDLSPELLLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL
-			, , , , , , , , , , , , , , , , , , ,			LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL GQGYVKRRDDRPQILMYAAPEPMGPCRGQA DVPFCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
						CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAIHEVPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS
475	1825	A	3754	1093	96	YDESENLDQETFFFGDEEEDEDEAYD GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR
476	1826	A	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	A	3763	267	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNITSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG PDEHPQDTDARDADGEAREREP/RRPSFAA*P VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA

COO III	LOPAIN	1 37.3	Labo	Predicted	D-45-4-4-3	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met hod	SEQ ID NO:		Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	nou	in	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide			location		I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN		corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino acid residue	
uence	1	1	914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
	i		ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
'	}	ŀ	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ì	1	peptide		/-possible nucleotide deletion, \-possible
	<u> </u>			sequence		nucleotide insertion
	J	<u> </u>	<u> </u>	<u> </u>		EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK
		1	1			DFFQKVSQVYVAIDERLASLKTDTFSKTREEK
ľ	1	i	ľ	{	1	MEDIFAQKEMEEGEFKNWIEKMQARLMSSS
1			İ		1	VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR
	-	1		İ		LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS
	ļ	1	1			AMDASPRNISPGLQNGEKEDRFLTTLSSQSST
ŀ	-	ĺ		1		SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE
ĺ	ŀ	1	i	1	ĺ	DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG
}		1				NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE
						PSSIIAFALSCKEYRNALEELSKATQWNSAEE
			1			GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE
	}	]	1	ļ		TEPOPTKKASGMLSFFRGTAGKSPDLSSOKRE
	1		i			TLRGADSAYYOVGOTGKEGTENOGVEPODE
	İ			1		VDGGDTQKKQLINPHVELQFSDANAKFYCRL
	<b>j</b>		1	i		YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ
[	1	1	1			ARGGKSGAAFYATEDDRFILKQMPRLEVQSF
	1				i	LDFAPHYFNYITNAVQQKRPTALAKILGVYRI
1				1 1		GYKNSONNTEKKLDLLVMENLFYGRKMAQ
	1	1		į		VFDLKGSLRNRNVKTDTGKESCDVVLLDENL
}	1		J			LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS
ļ			İ		ļ	HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD
		1				KKLEMVVKSTGILGGQG*MPTVVSPELYRTR
ŀ	1		l			7
100	1830	<del> </del>	3777	251	3	FCEAMDNYFLMVPDHCTGLGLNC
480	1830	Α	3///	231	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI
		1	1			KLNIKDPAITLDVYPNEVKNYVRTKTYTQMF
100	ļ	1	10000	1 222		I/ANFIMAKSWKQPTHPSVRT
481	1831	A	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID
	1	1	}	}		SIEANAESSEVLVERAPGQLQRPA\YYQKKSR
İ	1	l	1	1	1	KKMCLVVLVQTAILLICERIM*VVYTTKWSPPI
	<u> </u>			<u> </u>		VLPVSCFQGQKFN
482	1832	Α	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M
}	į.	ł	ł	ł	ł	PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA
	1		ļ	1		KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S
						/L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG
		1	1			KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR
1	1	1	ĺ	Í		PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK
	ļ	1	l		1	LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI
<b> </b>	1	1		Ī	1	QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\
1		1		1		SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS
1	[	1	1	1		PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP
		1	1	1		A*FFVFLVE\QGFTMLARMVSIS*PQ/CDLPAS
	1	1		1	1	VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC
	1	1	ì	1	1	SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG
1	Ì	1		{		SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS
	1	1		1	1	PDLVIRPPRPPKVLGLQA
485	1835	A	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF
403	1033	١^	3002	1.*	239	KOFSCLSLPSSWD*RVPTSRPAKF/CVIF*DGV
i	1		1	1	1	
400	1000	+	2011	270	00	SHCQPGWSAVVQPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN
	1	1				MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG
	1		1	ŀ	1	FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL
487	1837	A	3814	771	320	PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA
487	1837	A	3814	771	320	PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP
487	1837	A	3814	771	320	PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA

NO: of NO: of location contact entitle sequence per per per per per per per per per pe	670 70	- CTO TD	1364	Lono	T 90 91 1 1		· · · · · · · · · · · · · · · · · · ·
uence evide seq- uence USSN location correspondit sequence   914	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence USSN 1949 0 corresponding to list at any properties of peptide sequence peptide seque	_		noa	1			
sequence where the sequence peptide sequence peptide sequence peptide sequence seque							
uence alique of peptide sequence peptide					1		
amino seid recidite of peptide sequence peptide sequence peptide sequence peptide sequence per sequence sequence per sequence seq	•	uence					
residue of poptide   sequence	uence	į	l	914			
### ### ### ### #### #################			1	1			
Sequence   nineleotide insertion						sequence	
1838 A 3818   1   781			1	1			1
LCILAPCLAGSSYAGOULFKGCDVKTTFTV HVPCTSCAAIKKQTCFSGWIRELPDQTQDC YEVQLGGSMVSMSGCRRKCRKQVVQKACC GYWGSKCHECPGGAETPCNGHGTCLDGMD NGTCVCQENFRGSACQECQPTRRFGPDQSS CSCVHVCNCHGFRGDGSCLTCAGVTGPHCD QELPVWQELGFPQNNPRIRKAPNCKCLPG*B RNGLIATNPCRP RNGLIATNPCRP 489 1839 A 3822 934 669 FFFSSMERSYTRLECSGAISAHLRILGSSNS ASAPVAGTIGACHAQLEVTLVETOFHHV QDGLDLLANLMHPPRPPKVLOFQA GCQSCWPAWPRIRRRGPASGARIGRKAPV GLGRWQDGRLAFCYHRRAPFHAPVLSGA ASRPEASGDCRAGRETAMATLECLMKAFFSIS KSFQQQQQQQQQQQQQQQQQQQQQQQQQPPF PPPPPPQLPQPPPQAQPLLPQCPPPPPPPPPP GPAVAEPLHPRKELSSTKKDKVNHCLTIC ENIVAQSVRNSPEPGKLLGJAMELFLLCSDNS KSFQVQQQQQQQQQQQQQQQQQPPF GPAVAEPLHRRKELSSTKKDKVNHCLTIC ENIVAQSVRNSPEPGKLLGJAMELFLLSSDNS ESDVRWADECLNSVKALMDSNNPRLQLEI YKEIKKNGAPRSLRAALWRFAELAHLYRPQG CRYYLVNLLPCLTRTSKRPESVGPGLAAAV KIMASFGMFANDNEIKVLKAFIANLKSSSTT RRTAAGAVSICGHBAKELSTKKDKVNHCLTIC LLVPVEDEHSTLLILGGYLLTLRYLVPLLQQLY VKEIKKNGAPRSLRAALWRFAELAHLYRPQG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDTSLKAFFAELAHLYRPG KDT	400	1020	<u> </u>		<del></del>	401	
HVPCTSCAAIKKQTCPSGWLRELPPQGTIQDC  YEVQLGGSMYSMSCRRCKCKQVVQKACCI GYWGSRCHECPGGAETPCNGHOTCLDGMDS  NGTCVCQENFEGSACQECQDPNRFGPCQSS  CSCVHOVCNHGPRGDGSCLCFAGYTGPHCD  QELPVWQLGFPGNNFLRKAPNCKCLPG*H  RNGLIATTNPCRP  RNGLIATTNPCRP  RNGLIATTNPCRP  A3822 934 669 FFPSEMESRSVTRLECSGAISAHLRILGSSNSS  ASAS*VAGTIGACHHAQLIFVRLVETOFFHIVY  QDGLDLLANLMHPPRPFRVLOFQA  450 1840 A 3825 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPV  GLORGVODGNELRFCFYLRRAPFIAPVLSGA  ASRPEASODCRAGREGTAMATLEKLIKKAFESI  KSEQQQQQQQQQQQQQQQQQQQQQQQQQQPQPPPPPPPPPP	488	1838	, A	3818	] <sup>1</sup>	781	
YEVQLGGSMYSMSGCRRKCRKQVVÖKÄCC   GYWGSRCHECFGGAETPCNGHCIDGMDB    NGTCVCQENFRGSACQEQPPNFFGFDCS    CSCVHGVCNIGPGGGSCLCFAGYTOFPHCD    QELPVWQELGFPQNPRIRRAPNCKCLPG*H    RNGLIATTPNPCRP    ASPASSTSTTRLECSGAISAHILRILGSSNS*    ASSAVAGTIGACHHAQLIPVTTCFHHVC    ODGLDILATIMHEPPLPKVLGFQA    490   1840   A 3825   79   9748   GCQSCWPAWPRLRRGFASAGARLGRAPP    GCQSCWPAWPRLRRGFASAGARLGRAPP    GLYCNQDCGPLRCFTHLPGASAGARLGRAPP    GLYCNQDCGPQQQQQQQQQQQQQQPPF    PPPPPPQLQPPPQACPPLLPQPPPPPPPPP    GPAVAGEPHLRRKELSATKKDNHCLTIC   ENIVAQSVRNSPEFQKLIGIAMELFILLSDNSNL*PRLQLE    YKEIKKNGAPRSIRAALWRFAELAHLVSSNST    CRYLVNLIPCLITTSTKREEDSVALWINLLG   LLVPVEDEHSTILLIGVLITTSTKPERESVALWINLG   LLVPVEDEHSTILLIGVLITTSTKPERESVALWINLG   LLTANGGIGQITAAKEESGGRSRSGSIVEL   AGGSSCSPVLSRQKGKVLLGEEALEDDS    SERSDVSSALTASVEDEIGGLASAGSVSTP-   SAGHDIITEQPRSQHTLQADSVDLASCDLTSSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSSSAGSTSAGSTSTAGSTSSAGSTSSAGSTSSAGSTSSAGSTSAGSTSTAGSTSSAGSTSSAGSTSSAGSTSSAGSTSAGSTSTAGSTSSAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSTAGSTSSAGSTSAGSTSTAGSTSSAGSTSAGSTSTAGSTSSAGSTAGST				į			
SYWGSRCHECPGGAETPCNGHGTLDGMINGTOLOGN		}	1				
NGTCVCQENFRGSACQEQPNRFGPDGSC   CSCVMGVCHIGPRGDGSCLCFAGYTOPHCD   OELPVWQELGPPQNNPRLRKAPNCKCLPG*H   RNGLIATPHPCRP   RNGLIATPHPCRP   ASAS*VAGTGACHAQLFVFLVETGPHHVC   QDGLDLINIMIPPPRPKVLGPQA   490   1840   A 3825   79   9748   GCQSCWPWPRLRRRGPASAGARLGRKAN   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   ASRPEASGDCRAGRETAMATLERLMKAPPIAPVLSGA   CREVILVALPULT LIGHTSCHEDIAPPIAPVLSGA   CREVILVALPULT LIGHTSCHEDIAPPIAPVLSGA   CREVILVALPULT LIGHTSCHEDIAPPIAPVLSGA   KURASGORANDELEVILL LIGHTSCHEDIAPPIAPVLSGA   KURASGORANDELEVILL LIGHTSCHEDIAPPIAPVLSGA   KURASGORANDELEVILL LIGHTSCHEDIAPPIAPVLSGA   KURASGORANDELEVILL LIGHTSCHEDIAPPIAPVLSGA   CREVILVALPULT LIGHTSCHEDIAPPIAPVLSGA   AGGGSSCSPVLSRQCKCHLGEEALEDDS   ESRSDVSSALTASVCDEGGELAASSGVSTPP   SAGHDITEQPRSQRTILQADSVCHAPPIAPVLSGA   ATRICHAPPIAPVLSGAAPVL			1	İ	ļ		
CSCVHGVCHGFGGDSLCFAGYTOFENG GELPVWGELGFPQNNPRLRAPNCKCLPG+B RNGLIATTNPCRP RNGLIATTNPCRP 490 1840 A 3825 79 9748 GCGSCWPAWFRRRGRASAGRUGKKAPW GLGCSCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKKAPW GLGCCWPAWFRRRGRASAGRUGKAPW GLGCCWPAWFRRRGRASAGRUGKAPW GLGCCWPAWFRRRGRASAGRUGKCAPW GLGCCWPAWFRRRGRASAGRUGCWAPP PPPPPPPPQCQQQQQQQQQQQQQQQQQQQQQQQQQQ		Ì	j		1	1	
GELPYWOGELGFPQNNPRLRKAPNCKCLPG-HE RNGLIATIPNECRP   A839   A 3822   934   669   FFFSEMESRSTRLECSGAISAHLRILGSSNS;   ASAS-VAGTIGACHHAQLPYFLVETGFHNV   QDGLDLIAMMIPPRPRVLIGQA    490   1840   A 3825   79   9748   GCQSCWPWRRRRRGPASAGARLGRKAPI   GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ASRPEASGDCRAGRETAMATLEKLMKAFPIAPVLSGA    ESPAWASEDCLKVKKALMORICHTCCENTVCAMABELAHLYRPOO    GPAYLVALLPCLTTSKRPRESDCRAVLANIANICALLANIANICANIANICALLANIA		1		l	ļ		
RNGLIATPHPCRP		1					
489 1839 A 3822 934 669 FFFSEMSRSYTRLECSGAISAHRELLGSINSE ASASY-NAGTIGACHHAQLIFY-DETGFHHYC QDGLDLLANLMIHPPRPKVLGFQA  490 1840 A 3825 79 9748 GCQSCWPAWPRURRGFASAGARLGKAPW GLQRGVQDGRLRFCFYLRPAFFIAPVLSOA ASRFASGADCQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ			1	1	i		
ASAS*VAGTIGACHHAQLIFVFLVETGFHHVC QDGIDLIANIMIPPRPRVLIGPA  490 1840 A 3825 79 9748 GCQSCWPAWPRIRRGPASAGARIGRKAPV GLPGRVQDGRPLRFCYLIPRAPFIAYLSGA ASRPEASGDCRAGRETAMATLEKLMKAPFSI KSFQQQQQQQQQQQQQQQQQQPPI PPPPPPPQLPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVHICLTIC ENIVAQSVRNSPEFQKLLGIAMEFLLCSDDA ESDVRWAJDECLNKVIKALMDSNLFRLQLEI YKEIKKNGAPRSLRAALWRFAELAHLVRPQI CRPYLVNLLCHTRISKRPSVGETLAAAVY KIMASFGNFANDMEIKVLLKAFIANLKSSSFT RRTAAGSAVSICOHSRRTQYSWLEVILAQQV KDTSLKGSFGGVTKEMEVSYSLEVLAQQV KDTSLKGSFGGVTKEMEVSYSLEVLAQQV KDTSLKGSFGGVTKEMEVSYSLEVLAQQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSFGGVTKEMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTREMEVSTALDQV KDTSLKGSGGVTAPSTALDQV KDTSLKGSGGVTAPSTALDQV KDTSLKGSGGVTAPSTALDGV KDTSLKGGGGGTADAG GTCCSLKSRSGTPHOSTALDGV KDTSLKGGGGGTADAG GTCCSLKSRSGVTAPSTALDGV KDTSLKGGGGGTADAG GTCCSLKSRSGTPHOSTALTATGTGCCG ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFILLISATLFGCCC ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFILLISATLFGCCC ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFILLISATLFGCCC ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFILLISATLFGTCG VTMENNLSKVLAVSHELITSTTRALTFGCCC ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFILLISHTIKATLYGGG CREEPPGGASVFLSKGGSASAASRQSDT GPVTRINSSIGGFYHPSVLSHINPSSLSPIRK GEEKEPGGASVFLSKGGSASAASRQSDT GPVTRINSSIGGFYHPSVLSLINPPSLSPIRK GEEKEPGGASVFLSKGGSASAASRSDST GPVTTSKSSLGSFYHPSVLSHDVLAKTH NYKVTLDLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEEILGYHKGGSASAASRSDST GREEPFGGASVFLSKGGSASAASRSDST GREEPFGGASVFLSKGGSASAASRQSDT GREEPFGGASVFLSKGGSASAASRQSDT GREEPFGGASVFLSKGGSASAASRSDST GR	480	1930	<u> </u>	2822	024	660	
ODGLDLIANLMIHPPRPKYLGFQA	707	1009	1^	3022	734	1 009	
1840   A   3825   79   9748   GCGSCWPAWPILRRGGPASAGÄLGIKAYW GLFGRWQDGPLRFCFYLRRAPFIAPVLSGA ASRPASGDCRAGET AMATLEKLMKAFESI KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ		1	1	}	ł	1	
GLPGRVQDGRPLRFCFYLRPRAFAPLYLSGA ASRPASAGDCRAGRET AMATLEKLMKAFESI KSFQQQQQQQQQQQQQQQQQQQQQQPPP PPPPPPPPPPPP	490	1840	A	3825	79	9748	
ASRPEASODCRAGRETAMATLEKLMKAPESI KSFQCQQQQQQQQQQQQQQPPP PPPPPPPQLPQPPPQAQPLPQPPPPPPPPP GRAVAEEPLHRKKELSATKKORVHCLITIC ENIVAQSVRNSPEFQKLIGIAMELFILLCSDDA ESDVRWYADECLNKVIKALMDSNI-RICLQEL YKEIKKNGAPRSIRAA HYAFAELAHLVRPQB CRPYL-VNLLPCLTRTSKRPEESVQETLAAAVI KIMASFGRPANDMEILYLLAFIANI.KSSSTI RRTAAGSAVSICQHSRRTQYPYSWLLNVLLC LLVPVEDEHSTILLIG-VLLTURYL-VPLQQQV KOTSLKGSFGVTRKEMESSAEQL-VQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGGQLTAAKSGGSGSRSGSIVEL AGGSSCSPVLSRCQKGKVLLGEEFALEDDS ESRSDVSSSALTASVKDEISGGRSRSGSIVEL AGGSSCSPVLSRCQKGKVLLGEEFALEDDS ESRSDVSSSALTASVKDEISGGLAASSGVSTP SACHDITTEQPRSQHTLQADSVDLASCOLTSS ATDGDEEDLISHSSSQVSAVPSDPAMDLINDG TQASSPISDSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIQQPQDEDEATGIPDEASE FRNSSMALQQAHLLKMSHCRQFSDSSVIK VLRDEATEFGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRX SVKALALSCVGAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYDBIGDPQVGATALLC GTLICSLISRSRFHVGDWMGTTRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGQLQIDVILLRNSSVULVRTEI LETLABIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPVRHVAAASI RLVPKLFYKCDQQQADPVVAVARDQSSVYI KLLMHETQPPSHSSVSTITRIYRGYNLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPLSASDESI KSCTYGMATMILTLISSSWPDLDSSVIK LLMETQPPSHSSVSTITRIYRGYLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPLSASDESI KSCTYGMATMILTLISSAWPPLDI.SAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMYGQLFSTSPISRK GKEKEFGEQASVPLSPKKGSEASARQOST GVPTTSKSSSLGSFYHLSPYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEELICYLKSPYLKHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEELICYLKSPYLKHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEELICYLKSPYLKHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEELICYLKSFRFFMMATVC VQLLKTLFGTNLASGFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCMAPYTHITQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLKI	770	1070	1	3023	''	//70	
KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPP				Į.			-
PPPPPPGLPG/PPPAAOPLLPGOPPPPPPPPPP GPAVAGEPLHRPKKELSATKKDR WHCLTIC ENVAGSVRNSPEPGKLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRU, QLE YKEIKKNGAPRSLRALMFABLAHLVRPQB CRPYLVNLLPCLTRTSKRPESSVQETLAAAV KIMASFGNFANDNEIKVLLKAFIANLKSSSPT RRTAAGSAVSIQHBKVSPSABQLVQVYEL LLVPVEDEHSTLLILGVLLTLTRYLVPLLQQQV KDTSLKGSFGVTRKWFSYSABQLVQVYEL LLVPVEDEHSTLLILGVLLTLTRYLVPLLQQQV KDTSLKGSFGVTRKWFSYSABQLVQVYEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVEL AGGGSSCSPVLSRKQKKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPP SAGHDIITGQPRSQHTLQADSVDLASGLASSGVSTPP SAGHDIITGQPRSQHTLQADSVDLASGLASSGVSTPP SAGHDIITGQPRSQHTLQADSVDLASGLASSGVSTPP SAGHDIITGQPRSQHTLQADSVDLASCDLTSS ATDGDEEDLSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTLQADSVDLASCDLTSS ATDGDEEDLSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTLQADSVDLASCDLTSS ATDGDEEDLSHSSSQVAVPSDPAMDLNDG TQASSPISDSSQTTLQADSVAVLROPBSSIVLD GTIDNYLGLQIGQPDEDEATGLIPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVVDK VLRDEATFPGDQPDEMPCRIKGDIGGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRX SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPFEQTVSDLINYIDHGDPQVRGATALLC GTLICSILSRSRFHVGDWGGTITRITGTNTFSL ACTPLLKRTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSWLVKTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILLGDEDPRVRHVAAASI IRLVFKLFYKCDQGQADPVVAARDQSSVIVX VKLLMHETOPPSHRSVTITRYRGYNLLPSITD VTMENNI.SRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI LLAGNLLAASAPKSLRSSWASSEEANPAATK QEEVWPALGDRALVPWVEQLFSHLLKVNIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFVLPSYLKLHLVLKATH- NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEELICYSCSPSRPMMATVC VQQLLKTLFGTNLASQFGLSSNPSKSQGRA QRLGSSSVRPGLYNGVPSLYSLOGLSINPSKSQGRA QRLGSSSVRPGLYNGVSTGLKSVSVASGRA	i		1	1	1	l	
GPAVAEPLIRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILLCSDDA - ESDVRMYADECLNKVIKALMDSNLPRLQLEI YKEIKKNGAPRSLRAALWRFAELAHLVRPQG CRYPLVHLIPCLTRSKRPEBSVQETLAAAVI KIMASFGNFANDNEIKVILKAFIANLKSSSPT RRTAAGSAVSICQHSRRTQYFYSWLLNVLLC LLVPVEDEHSTLLILGVLTTLRYLVPLLQQQV KDTSLKGSFGVTKKEMEVSPSAEQLVQVYEI THEHTOHODHNVTTOAIBLQQLTRTPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVEI AGGGSSCSPVLSKKQKGKVLIGEEAALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSFISDSSQTITEQPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE, FRNSSMALQQAHLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDLNYIDHGDPQVRGATAILC GTLICSILSRSRPHVGDWMGTRTLTGGTPSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGIQLIDVLTLRNSSYWLVRTEI LETLAEDFFLVSPLEAKENHLRGAHNYTG LKI.QERVLNNVHILLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVRDQSSVYI KLIMHETQPPSHFSVTITRYRGYNLPSITD VTMENNLSRVLAAVSHELITSTTRALTFGCCE ALCLESTAPPCVIREGWHGVPPLSASDESI KSCTVGMATMILTLLSSAWPPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASSEEANPAATK QEEVWPALGDRALVPMVEQLFSHLIKVNIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRK GKEKEPGGASVPLSKKGSFASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHEVLKATH- NYKVTLDLQNSTEEFGGFLRSALDVLSQILEI  ATLQDIGKCVEELICYKGSFSRPMMATVC VQQLLKTLFGTNLASOPDGLSSNPSKSQGAA QRLGSSSVRPGLYHCFNAPTHFTQALAD SLRNMVQAEQENDTISGWFDVLQKVSTQLKI		1					
ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEI YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVI KIMASFGNFANDNEIKVLLKAFIANLKSSST RRTAAGSAVSICQHSRRTQYFYSWLLNVLLQ LLVPVEDEHSTLLILGVLLTRYVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYSEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAKEPSGGRSGSIVEL AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGSVTEL SAGHDITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSFISDSSQTTTEGPDSAVTPSDSSIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDL VIRDEATPEDQENKPCRIKGDIGGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRX SVALALALSCVGAAVALHPESFFSKLYKYVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATALLC GTLICSILSRSRPHVGDWGGTITLITGNTFSL ACPILLRKTLKDESSVTCKLACTAVRICVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVFKLFYKCDQGQADPVAVARDQSSVYI XLMHETQPSFHSKYVATARDSVILTSTRALTFGCCE ALCLLSTAFPCUWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWPPLDLSAHQDAI LLAGNLLAASAPKSLRSSWASSEEANPAATK QEEVWPALGDRALVPWVEQLSHILLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEEPGEQASVPLSPKKSGSEASAASRQSDT GPVTTSKSSSLGSFVLIPSYLKLHDVLKATH- NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI  ATLQDIGKCVEELICYKSCFSREPMMATVC VQQLLKTLFGTNLASQPGLSSNPSKSQGRA QRLGSSSVRPGLYHCYNLRCSLLATI			ľ	1	1		
ESDVRMVADECINĞVIKALMOSNIPRIQLEI YKEIKKNGAPRSLRAALWRFAELAHLVRPQI CRYYLVNLLPCLIKTISKRPESS VQETLAAAVI KIMASFGRFANDNEIKVLIKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILKAFIANLKSSIST RRTAAGSAVSICOHSVILTIPUT VPLLQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL LLVPVEDEHSTLLILGVLLTRYLVPLLQQV KDTSLKGSGFGVTRKEMEVSPSAEQLVQVYEL LAGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTF SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDLISHSSQVAVPSDPAMDLINDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEBATGLIPDEASE, FRNSSMALQQAHLIKNMSHCRQFSDSSVDK VLRDEATEPGDQENKPCRKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEQVYSDILNYUDHDPQVRGATAILC GTLICSILSRSRPHYGDWMGTIRTITONTPSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFALVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILLGEDPRVRHVAAASI RLVPKLFYKCDGQAPDVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIVRGYNLLPSITD VTMENNLSRVIAVSHELITSITRALTFGCCE ALCLLSTAFFVCUSGGAPDVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIVRGYNLLPSITD VTMENNLSRVIAVSHELITSITRALTFGCCE ALCLLSTAFFVCUSGGMPOVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIVRGYNLLPSITD VTMENNLSRVIAVSHELITSITRALTFGCCE ALCLLSTAFFVCUSLGWHCGVPISASDESI KSCTYGMATMILTLSSAWPPLDLSAHQDAI ILAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVARGPAIKAALPSHLPSTIRRLTFGCCE ALCLLSTAFFVCHSLGWHCGVPISASDESI GEVEVPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVARGPAIKAALPSHLPSTIRRLTFGCCE ALCLLSTAFFVCHSLGWHCGVPISASDESI GEVEVPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVARGPAIKAALPSHLPSTIRRLTFGCCE ALCLLSTAFFVCHSLGWHCGVPISASDESI GETTICSTERGEGFGAALDVLSQULEI ATLQDGKCVEEILIGYLKSCFSREPRMATVC VQQLLKTLFGTNLASGFGGLSSNPSKSQGRA QRLGSSSVRPGLYYYCFMAPYTHFTQALADA SLRRMVQARQENDTSGWPDVLQKVSTQLEI	Ì						
YKEIKKNGAPRSLRAALWRPAELAHLVRPOK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVI KIMASFGMFANDNEIKVLLKAFIANLKSSSFT RRTAAGSAVSICQHSRRTOYFYSWLLNVLLC LLVPVEDEHSTILILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELQQLFRTPPPEL LQTLTAVGGIGGLAAKEESGGRSRGSIVEL AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITEQPRSQNTLQASSVDLASCDLTSS ATDGDEEDILSHSSQVSAVPSDPAMDLNDG TQASSPISSQTTTEGDDSAVPSDPSSEIVLD GTDNQYLGLQIQQPQDEDEEATGILPDEASE, FRNSSMALQAHLLKNMSHCRQPSDSSVDK VLRDEATEFDGDENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRX SVKALALSCVGAAVALHPSSFFSKLYKVPLD TTEVPPEQQVYSDILNYIDHGDPQVRGATALLC GTLICSILSRSRPHVGDWMGTIRTLTGNTPSL ADCIPLLRKTLDESSVTCKLACTAVRNCVM SLSSSSYSELGQLIIDVLTLRNSSYWLVRTEL LETLAEIDFRLVSFLEAKARNLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVAVARDQSSVVY KLLMHETQPPSHFSVSTITRIVRGYNLLPSITD VTMENNLSRVLASVELLTSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNILAASAFKLSSWASEELERSTRATEGCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNILAASAFRLSSWASEELERSTRATEGCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNILAASAFVLSTSTRIVRGYNLLPSITD VTMENNLSRVLASRESSWASEELERSTRATEGCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLSSAWFPLDLSAHQDAI ILAGNILAASAFVLSPLSTSTRENTRATEGCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLSSAWFPLDLSAHQDAI ILAGNILAASAFVLSSWASEELRSTSTREEARAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAFGPAIKAALPSLITNPPSLSFIRRK GKEKEPGGQASVPLSPKKGSEASARGNOTH QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILIGYLKSCFSREPMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYTYSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYTYSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYTYSCFSREPFMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYTYSCFSREPFMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA	}		Į.	1	1		
CRPYLVNILPCLTRTSREPESVOETLAAAV KIMASFGNFANDNEIKVLLKAFIANLKSSSPT RRTAAGSAVSICQHSRRTQYFYSWLLNVLIC LLVPVEDEHSTLLLGVLLTRYLVPLLQQQ KDTSLKGSFGVTRKEMEVSPSAGDLVQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKESGGRSRSGSIVEL AGGGSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQVLGLQIQQPQDEDEEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRKGDIGGSTDDDS APLVHCVRLLSASFILTIGGKNVLVPDRDVRX SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEQYVSDLNYUDHDDPQVRGATALC GTLICSILSRSRPHYGDWMGTIRLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLOLIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEPRVRHVAAASI IRLVPKLFNKCDQGQADPVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCUWSLGWHCGVPFLSASDESI KSCTVGMATMILTLLSSAWPPLDLSAHQDAI ILAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHYLDDVAFGPAIKAALPSLTSPPSLSFRRK GKEKEPGGQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKHDVLKATH, NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMATVC VQQLLKTLFGTNLASGFDGLSSNPSKSQGRA QRLGSSSVPRGLJYTYCFMAPYTHFTQALAD, SLRRNMVQARQENDTSGWFDVLQKVSTQLKI	1			İ			YKEIKKNGAPRSLRAALWRFAELAHLVRPOK
KIMASSGNFANDNEIKVLIKAFIANILKSSSFT RRTAAGSAVSICQHSRRTQYFYSWLINVLLOL LLVPVDEHSTLLILGVLLTIRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGIVEL AGGGSSCSPVLSRKQKGKVLLGEEFALEDDS ESRSDVSSSAT TASVKDEISGELAASSGVSTP SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSISSSQVSAVPSDPAMDLNDG TQASSPISDSSOJTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEFATGILPDEASE, FRNSSMALQQAHLLKMSHCRQFSDSSVDK VLRDEATEPFGQENKPCRIKGDIGGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRA SVKALALSCVGAVAHPESFFSKLYKVPLD TTEYPEGQYVSDLNYDHGDPQVRGATAILC GTLCSILSRSRPHVGDWMGTRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLILRNSSYWLVRTEI LETLAEIDFRLVSFLEAKENLHRGAHHYTG LKLQFRVLINVHILLGDEPPVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLSSITD VTMENNLSRVIAAVSHELITSTTRALTGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWPLDLSAHQDAI ILAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGFAIKAALPSLTNPSLSPIRRK GKEKFFGGASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH, NYKVILDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEELGYLKSCFSREPMMATVC VQQLLKTLFGTNLASGPFGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD, SLRNMVQAEQENDTSGWFDVLQKVSTQLKI							CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
LLVPVEDEHSTLLILGVLLT.RYLVPLLQQU KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVEL AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDIITEQPRSOHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSGTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE, FRNSSMALQAHLLKNMSHCRQPSDSSVDK: VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHESFFSKLKYVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCEPLLRKLDESSVTCKLACTAVRICVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKILQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESS KSCTVGMATMILTLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPEGQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHL NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQQLLKTLFGTNLASOFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADJ SLRNMVQAAEQENDTSGWFDVLQKVSTQLKT			1		-		KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEEL TLHHTQHQDHNVVTGALELLQQLFRTPPEEL LQTLTAVGGIGQLTAAKEESGGRSRSGIVEL AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS EERSDVSSSALTASVDEISGELAASSGVSTF SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK; VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEGYVSDILNYDHGDPQVRGATAILC GTLICSILSTSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGQLIIDVLTLRNSSYWLVRTEI LETLAEDFRYSDILVSFLEAKAENLHRGAHTYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVFKLFYKCDQGADPVVAVARDQSSVYI KLLMHETOPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVLAAVSHELITSTTRALTFGCCE ALCLLSTAFFVCIWSLGWHGGVPPLSASDESI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSTRPVCIWSLGWHGGVPPLSASDESI KSCTVGMATMILTLLSSTRPVCIWSLGWHGGVPFLSASDESI KSCTVGMATMILTLLSSAWPLDLSADPSI KSCTVGMATMILTLLSSTRPVLINGVSLGHEI ATLQDIGKCVEELIGYLKSCFSREPMMATVC QEEVWPALGDRALVPMVEQLFSHLLKVINGC AHVLDDVAPGPAIKAALPSILTNPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYLPSVLKLHDVLKATHL NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD, SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1					İ	RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
TLHITQHODHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRGSIVGSIVEL AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITGPRSQHTLQADSVDLASCDLTSS ATDGDEEDIJSHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIQQPQDEDEEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK: VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEGQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAFQPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHL NYKVTLDLQNSTEKFGGFLRSALDVLKGTHL NYKVTLDLQNSTEKFGGFLRSALDVLKSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYHPSYLKLHDVLKATHL NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC			i			1	LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
LQTLTAVGGIGQLTAAKEESGGRSRSGSIVEL AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPV SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDLISHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEGQYVSDILNYUDHGDPQVRGATAILC GTLICSILSRSRPHVGDWMGTIRTLTGNTFSL ADCTPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGADPVVAVARQQSSVYI KLLMHETQPPSHFSVSTITRIVRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSVLKLHDVLKATHJ NYKVTLDLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASGFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADJ	1		1	1			KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
AĞGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVRDEISGELAASSGVSTPC SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTGGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDK. VLRDEATEFGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAIL.C GTLICSILSRSRPHVGDWMGTIRTLTGNTTSL ADCIPLIKKTLKDESSYTCKLACTAVRNCVM SLCSSSYSELGQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLLGDEDPRVRHVAAASI IRLVPKLTYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHPSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGFAIKAALPSLTNPPSLSPTRK GKEKEPGEQASVPLSPKKGSEASAASROSDT GPVTTSKSSSLGSFYHLPSYLKHDVLKATH, NYKVTLDLQNSTEKFGGFLRSALDVLSQLEEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLPGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA			ŀ	1	1		TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
ESRSDVSSSALTASVKDEISGELAASSGVSTPC SAGHDIITEGPRSQHTILQADSVDLASCDLTSS ATDGDEDII.SHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTILCSILSRSRPHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGGRASVPLSFKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH, NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGITNLASGFFGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI						1	LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK: VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFFVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMYEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSFKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH, NYKVITDJLQNSTEKFGGFLRSALDVLSQLEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADJ SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	}		j	1	j	1	AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGADPVVAVARDQSSVTV KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTILSSAWFPLDLSAHQDAI ILAGANLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH- NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFFGLSSNPSKSQGRA QRLGSSSVRPGLYHVCFMAPYTHETQALADJ SLRNMVQAEQENDTSGWFDVLQKVSTQLKI							ESRSDVSSSALTASVKDEISGELAASSGVSTPG
TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIQQPQDEDERATGILPDEASE FRNSSMALQQAHLLKNMSHCRQPSDSSVDK: VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTINPFSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSVLKLHDVLKATHL NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADV SLRNMVQAEQENDTSGWFDVLQKVSTQLKI	1						SAGHDIITEQPRSQHTLQADSVDLASCDLTSS
GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQFSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAIL.C GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTILSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSVLKLHDVLKATHL NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFFDLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADV SLRNMVQAEQENDTSGWFDVLQKVSTQLKI		1	I				h
FRNSSMALQQAHLLKNMSHCRQPSDSSVDK: VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDDVRV, SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRPHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG: LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSKPIRKK GKEKPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI			1				
VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GRVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREFMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT					1		
APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILLGDEDPRVRHVAAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKPGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI		ł	ì	1	· ·	1	
SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRPHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSVLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI				1	1		
TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT		1	1	1		1	
GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASJ IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEII ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	1	ļ	1		,	•
ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG; LKLQERVLNNVVHILLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT: GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH, NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT		[	1	i		1	
SLCSSYSELGLQLIIDVLTLRNSSYWLVRTEI LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEII ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT			1	1			
LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILITLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFIRSALDVLSQILEII ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT			1			1	
LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIO AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	1					
IRLVPKLFYKCDQGQADPVVAVARDQSSVYI KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIO AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	1			1	1	
KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIO AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT			1	1			
VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT			1				KLLMHETOPPSHFSVSTITRIYRGYNLLPSITD
ALCLLSTAFPVCIWSLGWHCGVPPLSASDESI KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT: GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1						VTMENNLSRVIAAVSHELITSTTRALTFGCCE
KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT: GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH, NYKVTLDLQNSTEKFGGFLRSALDVLSQILEII ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT					1		ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT: GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFIRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1						KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	1					ILAGNLLAASAPKSLRSSWASEEEANPAATK
GKEKEPGEQASVPLSPKKGSEASAASRQSDT: GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT							QEEVWPALGDRALVPMVEQLFSHLLKVINIC
GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI		1	1				AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
NYKVTLDLQNSTEKFGGFLRSALDVLSQILEI ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI		}	}				GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKI							GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT							NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1		1				ATLQDIGKCVEEILGYLKSCFSREPMMATVC
SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	]	İ	1	1	1		VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
	1		1	1	1		QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
, , , , , , , , , , , , , , , , , , ,		1	1		1		
	L	<u> </u>		<u></u>	<u>l</u>	<u> </u>	NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGAALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
			*			PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY
						LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
					,	INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATIGALISHEKLLLQINPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE ADAPAPSSPPTSPVNSRKHRAGVDIHSCSOFL
						LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGVLYVLECDLLDDTAKQL IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS
						EESTPSIIYHCALRGLERLLLSEQLSRLDAESL VKLSVDRVNVHSPHRAMAALGLMLTCMYT GKEKVSPGRTSDPNPAAPDSESVIVAMERVS VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA TWSLSCFFVSASTSPWVAAILPHVISRMGKLE QVDVNLFCLVATDFYRHOIEEELDRRAFQSV
491	1841	A	3826	469	302	LEVVAAPGSPYHRLLTCLRNVHKVTTC SNPPASASRVAGITGVHQHAWLIFVFLVEMEF
				392	88	HHVGQAVLKLLISGDLPVSASQSA VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE
492	1842	A	3836	372	00	FQSEWTAVV/P/EFTATQSEVADWFKDMQVP SVPIQQFPTEDWST*PTMNDWSATSTAQTTE WVRITTEWP

			T 45.4	- P	F 70 - 21 - 1 - 1 - 1	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted beginning	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq- uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uclice	İ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
nence	ŀ	ļ	114	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ì	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Į	1	peptide		/=possible nucleotide deletion, \=possible
	1	1		sequence		nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
4,7,5		,	1	1		KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK
	1	1	İ			CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL
		}	ł			VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL/ELGLLQAPPPGF
		}		į	ł	KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG
		1		ł	1	FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
					<u> </u>	ARPQDIDFLYAHQGRCWFRLL
495	1845	A	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL
	Į.	ļ				WADKYRPRKPRFFNRVHTGFEWNKYNQTHY
		1	İ		1	DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL
		1		1	1	EACADNKDFAILRFHAGPPYEDIAFKIVNREW
		1			ì	EYSHRHGFRCQFANGIFQLWFHFKRYRYRR* RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
		1	ì	1	Į	OGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC
l			1	ļ	· I	HGELRRHWDRLA*GPDATEGALGASFEHEG
	ŀ	1	1	1	Ì	GQQPPADLTVQADTLHRPSARLGGAHRACPK
1	ł					RRPHRVLWRWARGAWAWRCQAREKQETQG
1	1		1	1		QPCHITGHPLGREAEPAAAGAAPALAHRPPF
		1				ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
) '					1	WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
	}	]	.		1	VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE
ì		1	1			KGEGLPPGIPPPGLLTGPW\SMRPVTPSFAHIR
		1		ŀ	ł	TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR
}				1		LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
ł	1	1	-			SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
ļ		1		1	İ	SFVLMELAYWQDRMFF
496	1846	A	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
				1	1.	LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG
		1	İ		i.	LKLLTSSALPALASQSAEITGMSHRIWPLPLLR
]	j	1	1	}	1	RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	A	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR
				İ		LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS
]	}	}		1	ł	PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR
1		1				RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ
					1	VPIL
498	1848	Α	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP
						LPCLANF\*FLVETGFHHVGQADLKLLTSGDP
j	'	)	ļ			PTSASESAGITGVSHRAWPRIHFLYWKTFFL
<u></u>				422	262	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI
499	1849	A	3863	423	263	KIGINLTKEVKYLYTENYITLMKEIK/DTDKW
1		1				KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP
		ł	ì	•	1	MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S
					1	FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI
		1	l l			LVGIALNL
600	1050	+-	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP
500	1850	Α	3803	12	15240	DLRPWASDLDIMGDAEGEDEVQFLRTDDEV
				Į.	Į	VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP
				1	1	TSNAQNVPPDLAICCFVLEQSLSVRALQEML
1		1			}	ANTVEAGVESSOGGGHRTLLYGHAILLRHAH
1	ļ	1		}	}	SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE
1	ĺ	1	1		1	ACWWTMHPASKQRSEGEKVRVGDDIILVSVS
	)	]		1	1	SERYLHLSTASGELQVDASFMQTLWNMNPIC
		1	- {		1	SRCEEGFVTGGHVLRLFHGHMDECLTISPADS
j	1	)		1	1	DDORRLVYYEGGAVCTHARSLWRLEPLRIS
1		1		Į.	1	WSGSHLRWGQPLRVRHVTTGQYLALTEDQG
			1	ļ	Į.	LVVVDASKAHTKATSFCFRISKEKLDVAPKR
	1	1	1	1		DVEGMGPPEIKYGESLCFVQHVASGLWLTYA
	i	1	4			D. Downer : But I of Date : date : date :

PCT/US01/03800

070 170	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID					nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning		F-Phenylalanine, G-Glycine, H-Histidine,
nucl-	peptide		in	nucleotide	location	
cotide	seq-	1 1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M-Methionine, N-Asparagine, P-Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		i .		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ		1		peptide	} .	/=possible nucleotide deletion, \=possible
1				sequence	•	nucleotide insertion
				Soquence		APDPKALRLGVLKKKAMLHQEGHMDDALSL
ì	ł	'				TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
	ļ	}			)	TROUGESQUARRIESTIVOLITAGE ESCOPE
ĺ	1	l .				GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
<b>.</b>	l	1	,			EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV
1	,		1	ł		LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
	Ì	1	1	ļ		VNLLYELLASLIRGNRSNCALFSTNLDWLVS
ĺ	ļ		İ	1		KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
l			1		ļ	KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV
1	1	1				RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
}	i	}	]		ļ	IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
1	]		1		}	ATHLRVGWALTEGYTPYPGAGEGWGGNGV
Į	1	1				CDDI VENCEDCI DI MACCINI PONTESCONI
1	1		1	Į.	1,	GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
1	1	1	1	i	1	LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
1			1			NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
İ		1	[	1	1	LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
Į.	İ		1	1		RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
İ	1	1	ł	ł		LERIREKLAENIHELWALTRIEQGWTYGPVRD
1.	1	1		1	1	DNKRLHPCLVDFHSLPEPERNYNLQMSGETL
1	ĺ	{	ŀ	1		KTLLALGCHVGMADEKAEDNLKKTKLPKTY
1	1.	i i		}		MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
	1	1	l.			NGHNVWARDRVGQGWSYSAVQDIPARRNPR
	1	1	1	i		LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
1	ŀ	1	i			NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
1		İ	1			NIEPPDQEPSQ VENQSKCDK VKITKAEKS I I V
	Į.	1	1			QSGRWYFEFEAVTTGEMRVGWARPELRPDV
i i		1	1		1	ELGADELAYVFNGHRGQRWHLGSEPFGRPW
-	1		1	Į.		QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
1		1	1	İ		ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
i	1		ì	1	1	VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
1		1	1		1	KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
1			1			LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
1	1	Ì	1		1	CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
1		1	1		Ì	· NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
1	1	]			1	AGGEAQPARAENEKDATTEKNKKRGFLFKA
[	-		1			KKVAMMTQPPATPTLPRLPHDVVPADNRDD
1	1		1	1		
	1	1	1			PEILLNTTTYYYSVRVFAGQEPSCVWAGWVT
}	1	1	1		1	PDYHQHDMSFDLSKVRVVTVTMGDEQGNV
1						HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
1	l		1		ľ	VIGCLVDLATGLMTFTANGKESNTFFQVEPN
1		1		j'	1	TKĪFPAVFVLPTHQNVIQFELGKQKNIMPLSA
	1			1		AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
		1		1	,	MPNHFLOVETRRAGERLGWAVQCQEPLTMM
1		į	1	1	1	ALHIPEENRCMDILELSERLDLQRFHSHTLRL
1		i	1	1		YRAVCALGNNRVAHALCSHVDQAQLLHALE
		ļ	1	1		DAHLPGPLRAGYYDLLISIHLESACRSRRSML
	1	1	ļ	1	,	SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
		[		1	i	CUCLETTE I DEBUTECEDOCENA AT DA ACA ADAD
1	1		1	1	ļ · .	GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
1	1	1	1	Į.	1	ARLSPAIPLEALRDKALRMLGEAVRDGGQHA
İ	İ	1	i	1	}	RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
	1	1		l	1	DVKQILKMIEPEVFTEEEEEEDEEEGEEEDEE
			1	1	1	EKEEDEETAQEKEDEEKEEEEAAEGEKEEG
1		1		1	1	LEEGLLOMKLPESVKLQMCHLLEYFCDQELQ
		1		1	1	HRVESLAAFAERYVDKLQANQRSRYGLLIKA
1		1		1	1	FSMTAAETARRTREFRSPPQEQINMLLQFKDG
		1	1	1	1	TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
				1	1	GEEEEPEEETTLGSRLMSLLEKVRLVKKKEEK
	1		J	J	1	DEPEREUR LEGANDOL OF LIGHT AUTOMACEA
		1				PEEERSAEESKPRSLQELVSHMVVRWAQEDF
1	1			1		VQSPELVRAMFSLLHRQYDGLGELLRALPRA
			1		İ	YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
]	}	)	1	}		ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
[	1		1	ļ		TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL
L						

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
l		1		peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
			<del></del>	sequence		CYFCRISRONORSMFDHLSYLLENSGIGLGM
						QGSTPLDVAAASVIDNNELALALQEQDLEKV
		1				VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
	,	]			1	ERYLDFLRFAVFVNGESVEENANVVVRLLIR
					ł	KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
				Į		DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
•					1	FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
					·	LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
	İ	1				MSASFVPDHKASMVLFLDRVYGIENQDFLLH
					1	VLDVGFLPDMRAAASLDTATFSTTEMALAV
				}	i	NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
1						MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
		]	]	j	]	TNHYERCWKYYCLPTGWANFGVTSEEELHL
	-		1			TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
1				1	1	AIAGALPPDYVDASYSSKAEKKATVDAEGNF
	į.	-		1		DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF
						DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
1	ļ	l		1		KDKETYRWPIKESLKAMIAWEWTIEKAREGE
1		<u> </u>		]		EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
		1		i		SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
	ļ	ļ				ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
						QELLKFLQMNGYAVTRGLKDMELDSSSIEKR FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
1						EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
		l	İ		1	TPAKVLGSGGHASNKEKEMITSLFCKLAALV
					ļ	RHRVSLFGTDAPAVVNCLHILARSLDARTVM
	}	1				KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
					ĺ	KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
		]			1	QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
	-		İ			TTKNTYVEKLRPALGECLARLAAAMPVAFLE
						PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
				İ	1	CPDIPVLERLMADIGGLAESGARYTEMPHVIE
1 .				l		ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM
		ł	ļ			KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
					1	AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
					1	VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
		İ			1	AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
-				İ		NEINNMSFLTADNKSKMAKAGDIQSGGSDQE
				1	1	RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN
	'	1	1			MCAPTDQDLITLAKTRYALKDTDEEVREFLH
1						NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
	1.			1		KKAVWHKLLSKORRRAVVACFRMTPLYNLP
	1			İ		THRACNMFLESYKAAWILTEDHSFEDRMIDD
1 .	1 .	1		1	1	LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT
				1		ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
	1			1		GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
		1				HTRGAAEMVLQMISACKGETGAMVSSTLKL
	1 .	1 .		İ		GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
	1	1				IQALMQTCSVLDLNAFERQNKAEGLGMVNE
					1	DGTVINRQNGEKVMADDEFTQDLFRFLQLLC
		1				EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLOESISDFYWYYSGKDVIEEOGKRNFSKAM
		1	1			SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
ł		}	}			DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
						LDLQKDMVVMLLSLLEGNVVNGMIARQMV
		1	1			DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
1						QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI
	<del></del>		٠			

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
			·		·	QFLLSCSEADENEMINCEEFANRFQEPARDIG FNVAVLLTNLSEHVPHDPRLHNFLELAESILE YFRPYLGRIEIMGASSRIERTYFEISETNRAQW EMPQVKESKRQFIFDVVNEGGEAEKMELFVS FCEDTIFEMQIAAQISEPEGPETDEDEGAGA AEAGAEGAEGAAGLEGTAATAAAGATARV VAAAGRALRGLSYRSLRRRVRRLRRLTAREA ATAVAALLWAAVTRAGAAGAGAAAGALGL LWGSLFGGGLVEGAKKVTVTELLAGMPDPT SDEVHGEOPAGPGGDADGEGASEGAGDAAE GAGDEEEAVHEAGPGGADGAVAVTDGGPFR PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG VDGVEEELPPBPEPEPEPELEPEKADAENGEK EEVPEPTPEPPKKQAPPSPPPKEEAGGEFWG ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS GGSSGWGLGAGEEAEGDEDENMVYYFLEES TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP LVIFKREKELARKLEFDGLYTTEQPEDDDVKG QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG DIYGRERIAELLGMDLATLEITAHNERKPNPP PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
501	1851	A	3869	467	665	HNLANYMFFLMYLINKDETEHTGQESYVWK MYQERCWDFFPAGDCFRKQYEDQLS VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
						LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	A	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD YRHAP\PLLTNF\*FLVEMGFCYVGQAGRKLL ASSDQSALASQSAGITGISTAPGPPFFFLNFEA GSCSVAQAGVQ
503	1853	A	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR TQKHTIYLIPYQVIFWSTGKDAMRSFMMPFY QKEYYENQ*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG NENTKLELRKVPPELNNISKLNEHFSRFGTLV NLQVAYNGDPEGALIQFATYEEAKKAISSTEA VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS ASSDLPQVLSTVLLA*QKQCIIQLL/WKAAQKT LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMQAGEEVTEL RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA VHGRGRGRGRGRGRGVPGHAVVDHRPRALEIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA VITFKTRAEAEAAAVHGARFKGQDLKLAWN KPVTNISAVETEEVEPDEEEQREIIIA  DAELSGTLSLVLTQCCKRIKDTVQKLASDHK
						DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLHRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLVYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEVDLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKQRDKRNRHLGR
508	1858	. A	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKJKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRMY\SP MAALILVADNAGGSHASKDANQVHSTTRRN SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHHHTWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPLFHHGL*AWLWCPELLLQGARRH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PPCHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRY\PSHLAESCISSRDRIP PSRPDRSRNSNSLSR
513	1863	A	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTTDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid; E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
İ				sequence		nucleotide insertion
				Soquenee	<del></del>	SLASSTVGLAGQVVHTETTEVVLTADPVTGF
4	l i				ì	GIQLQGSVFATETLSSPPLISYIEADSPAERCG
				•	<b>{</b>	VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI
				ļ		TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN
						VELGITISSPSSRKPGDPLVISDIKKGSVAHRT
}	ł	l		1	1	GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC
	ł	ł				EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR
	ļ	l		ł	1	YGGPLG\ITISGTEEP\FDL*IISSLTKGGLAERT
1	Ì	1			1	GAIHIGDRILAINSSSLKGKPLSEAIHLLQMAG
		l			Į.	ETVTLKIKKOTDAQSASSPKKFPISSHLSDLGD
		1	)		}	
! '	1	1		Ì		VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD
				<b>\</b>		SWDGSA\IDTS\YGTEGT\SFQASGY\NFNTYD
	-	1				WRSPKQRGS\LSPVT\KPRSQTYPDVGLSYED
i	1	ł	1			WDRSTASGFAGAA\DSAETEQEENFWSQALE
[		1		}	}	DLETCGQSGILRELEATIMSGSTMSLNHEAPT
[		1	1			PRSPAGSDRPSFQERSSSRPHYSQTTRSNTLPS
· '	1	l				DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT
Į	i				1	LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA
						GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV
	ł	1				VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG
•	ł	(				D*SEQNSAFFQQPSHGGNLETREPTNTL
514	1864	Α	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSLAPANG
1	1					NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH
ł	1	ł			1	SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI
İ	ļ		ļ		1	LSEFCMELTGIKQAQVDEGVPLKICLSQFCK
ì		1	Ì			WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL
	]	ļ	]			VR*RISYTY*SKHKSKGC
515	1865	A	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL
1 3.2	1.002				1	DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC
	İ	1				PNFILEEGTDLIF\*QVKHNPCHRLTPEEGTVQL
		l		1	1	NRADS
516	1866	A	3977	2	1357	KMLC/OKESNYIRLKRAKMDKSMFVKIKTLGI
1 3,10	1800	1	3577	1	1337	GAFGEVCLARKVDTKALYATKTLRKKDVLL
1		ì	1			RNQVAHVKAERDILAEADNEWVVRLYYSFQ
	1					DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL
i	1		1	1		ARFYIAELTCAVESVHKMGFIHRDIKPDNILID
1		1		1	1	RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP
	1	1	İ	1	-	RODSMOFSNEWGDPSSCRCGDRLKPLERRAA
		1	1		1	
		1		1	1	RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL
			1	1	1	CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWOTSLHIPPQAKLSPEASDLIIKLCRGPE
		1				DRLGKNGADEIKAHPIF*NOFDFSO*PEDSRS
	i	ļ				
	l		1		1	AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN
						EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF
}	}	}	1		1	FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD
						QNTGSEIKNRDLVYV
517			3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY
1	1867	A	1		1	VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD
1	1867	A	1		1	
	1867	A			1	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA
	1867	A				*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518	1867	A	3986	974	666	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA
518			3986	974	666	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518			3986	974	666	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F
518			3986	974	666	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH
	1868	A				*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF
518			3986	974	126	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF NQGLRHVGLCRTCLVNQMFASSILGKSHHHS
	1868	A				*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV  SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF  NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC
	1868	A				*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS
	1868	A				*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV  SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF  NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC

NO: of nucleotide conide sequence	0							
Deathing	=Cysteine,	Amino acid sequence (A=Alanine C=Cysteine	Predicted end	Predicted	SEQ	Met	SEQ ID	SEQ ID
Seq-	) _=!d!	D=Aspartic Acid, E=Glutamic Acid,				hod		NO: of
Sequence	stidine,	F=Phenylalanme, G=Glycine, H=Histidine,		*****			peptide	nucl-
1870   1870   1871   1871   1871   1871   1872   1874   1874   1874   1874   1874   1874   1874   1875   1875   1875   1877	) )1!	I=Isoleucine, K=Lysine, L=Leucine,					seq-	cotide
mino acid residue of peptide sequence public sequence public sequence   T=Threonine, X=Unknown, *=Stop codo	roline,	M=Methionine, N=Asparagine, P=Proline,			09/496		uence	seq-
residue of peptide sequence	e,	Q=Glutamine, R=Arginine, S=Serine,			914		]	uence
peptide sequence	pnan,	T=Threonine, V=Valine, W=1 ryptophan,			]		1	
Sequence	odon,	Y=Tyrosine, X=Unknown, *=Stop codon,	sequence		'			
LRCLGGEKHKSGLHARPVIVPSLELI	ssible			peptide			<u> </u>	
AHVPADLLITTLPSYYPFC				sequence				
520	LELHYDMDSI					•		
SIALLPRLYSNSWPQAILPPRPPKVL		AHV\FADLLLIITLPSYYIPFC			1	1		
SIALLPRVSNSWPQAILPPRPPKVL	F*T\FFCRDR/	QSFRLSLLSSWDYRHM*PRLANF*T\FFC	698	882	3999	A	1870	520
1871	KVLGLQT	SLALLPRLVSNSWPQAILPPRPPKVLGLC						
PPTSASHVAGATGTHHHAWLLSV	LQPPSPG\SSN	FFF*ETVSCSAS*AGVRSHDNSSLQPPSP(	1178	1346	4011	A	1871	521
S23	SV	PPTSASHVAGATGTHHHAWLLSV				l	1	
S23	LRTRPLEFAA	QGIALLTRMGESVKHVTGGYKLRTRPLI	377	2	4015	A	1872	522
1873   A   4018   341   19   ERVIENDIÇ QAQREPHIFNARRISPE	<b>EEIEYLVELR</b>	IGDYLDTFALKLGTIDRIAQRIIKEEIEYL	[	-	10.2	1	10,2	J=2
S23	VSACIGNCST	EYGPVYSTWSALEGELAEPLEGVSACIG					1	
1873   A   4018   341   19		AL*ELTDDMTEDFLFVLREYILYSDSMK		ļ		ļ	j	
KVKEVCKTSKS/GQVIYKGVSIRLR/ L*NRREWDEARKVLKEKQVFLSKMV GNEGDITSFPAK   SPILRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGVKWCNLGSLA    FFLRWSL/DSVAQAGWWHD    FFLRWSL/DSVAQAGWQWHD    FFLRWSL/DSVAQAGWQWHD    PPRLKQ/F/SHLSPPSIWDYRRVPPCLA    VETGSCQPCLQLLGSSNPPASASQS,   QQPE*SFDIRFACVIAALRETFQCL    VETGSCQPCLQLLGSSNPPASASQS,   QQPE*SFDIRFACVIAALRETFQCL    NKIINRPTHPVESSF    RSSAEDGWKADKP/VDG*TPGEDH    LHHSSESQLHHSVKSPPSLSFRLM    DVAVYFTTKEWAIMGPAERALYR    S27			19	341	4018	<u> </u>	1973	522
L*NRREWDEAIKVLKEKQ\FLSKMV   GNEGDITSFPAK	LRANFLAEP	KVKEVCKTSKS/GOVIYKGVSIRLRANF)		1	1010	\ ^	1073	323
S24	KMVYPANLSF	L*NRREWDEAIKVLKEKO\FLSKMVYPA	l	}	}	}	1	
1874			ļ				1	-
TFFSCLSLPSSWDYRHPPPRLAN*LI RQGFTVLARMVLIS*PHDLPASASQ SHCSWPTSSILS   SHCSWPTSSILS	GSLOAPPPGF	FET RWSL/DSVAOAGVKWCNLGSLOAP	7/13	1067	4020	<del> </del>	1974	504
RQGFTVLARMVLIS*PHDLPASASQ   SHCSWPTSSILS	V*LTNFLCF**	TPESCLSLPSSWDYRHPPPRLAN*LTNFI	,43	1007	4020	^	10/4	324
SHCSWPTSSILS	ASOS AGITGI.	ROGETVI ARMVI IS*PHDI PASASOSAC					l	]
525	(0					1	1	
PPRLKQ/F/SHLSPPSIWDYRRVPPCI	IGO LIZHOHV		251	701	4001	<del> </del>	1000	
VETGSCQPCLQLLGSSNPPASASQS. QGQPE*SFDIRFACVIAALRETFQCL NKIINRPTHPVESSF  526 1876 A 4024 80 341 TPSSTSRGTEQQSSKMAWQRREEL RSSAEDGWKADKP/VDG*TPGEDH LHIHSSESQLHHSVKSPPSLSFRLM  527 1877 A 4026 593 230 DFYLYPERKRGQMMTAVSLTTRI DVAVYFTTKEWAIMG\PAERALYR YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT  528 1878 A 4028 1160 242 GTSELLCIQRWNWGPAFPPRPGLAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDFETPRSS\GSN NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\GA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	PCI VNESIEE	DDDI KU/EKRI EDDEIMDABBABUI ANI	331	/81	4021	A	1875	525
QGQPE*SFDIRFACVIAALRETFQCL   NKIINRPTHPVESSF	SUSAGIAGISH	VETGECOPCI OLI GESNIPPASASOSAGI.		1	1	ł		i
NKIINRPTHPVESSF	OCT CGA GB VO	OCODE#CEDIDE#CVIIAAI PETEOCI CSA		Ì	}		l	
526 1876 A 4024 80 341 TPSSTSRGTEEQQSSKMAWQRREEL RSSAEDGWKADKP/VDG*TPGEDH LHIHSSESQLHHSVKSPPSLSFRLM 527 1877 A 4026 593 230 DFYLYPERKRGQMMTAVSLTTRI DVAVYFTTKEWAIMGVPAERALYR YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT GTSELLCIQRWNWGPAFPPRPGILAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSDKPSRDPETPRSSIGSW NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\LGGA\LGTGR\LI WEQGQD\HDKENQHFPLVES 529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	QCLC0A5K VIV					-	Ĭ	•
RSSAEDGWKADKP/VDG*TPGEDHI LHIHSSESQLHHSVKSPPSLSFRLM  527 1877 A 4026 593 230 DFYLYPERKRGQMMTAVSLTTRI DVAVYFTTKEWAIMG\PAERALYR YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT  528 1878 A 4028 1160 242 GTSELLCIQRWNWGPAFPPRPGLAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETTPVASQSDKPSRDPETPRSS\GSN NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\LIGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	OFFICIAL NIVE	TENEGRAP CARROCCENTY MODDEEN EN	1 211			<u> </u>		L
527 1877 A 4026 593 230 DFYLYPERKKRGQMMTAVSLTTRI DVAVYFTTKEWAIMGPAERALYR YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT GTSSLSRNDWRAGWIGYLELRRYT GTSSLSRNDWRAGWIGYLELRRYT VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPP WNQTEFPSKQVFSK ETPVASQSDKPSRDPETPRSSIGSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\UGA\UGGR\LI GTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	YEEVEUTIAAY	PECATOCAL ADVIDADOS TROCTOR DEL	341	80	4024	A	1876	526
527 1877 A 4026 593 230 DFYLYPERKKRQMMTAVSLTTRI DVAVYFTTKEWAIMG\PAERALYR YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT GTSELLCIQRWNWGPAFPPRPGLAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPP WNQTEFPSKQVFSK ETPVASQSDKPSRDFETPRSS\GSW NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	IM	K22VEDA MVVDVL ADQ. LACEDUPL II			1	1		
DVAVYFTIKEWAİMG\PAERALYR YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT  TGSSLSRNDWRAGWIGYLELRRYT  VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDFETPRSS\GSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  TSP  1879  A 4039  A 4039  B TGSSLSRNDWRAGWIGYLELRRYT  VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL  VFETEDSKSNLPPEVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDFETPRSS\GSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  TSP  TSP  TSP  TSP  TGSSLSRNDWRAGWIGYLELRRYT  TGSSLSRNDWRAGWIGYLELRRYT  VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL  USUMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL  TGSSLSRNDWRAGWIGYLELRYT  VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLELRYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGYLERYT  TGSSLSRNDWRAGWIGHT  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWRAGWIGH  TGSSLSRNDWR						<u> </u>		
YGGCGPL*CHPTSKPALVFS\LEQGI TGSSLSRNDWRAGWIGYLELRRYT  528 1878 A 4028 1160 242 GTSELLCIQRWNWGPAFPPRPGLAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSSGSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	INPUESVAFE	DFYLYPERKKRGQMMIAVSLITRPQES	230	593	4026	Α	1877	527
TGSSLSRNDWRAGWIGYLELRRYT  528  1878  A 4028  1160  242  GTSELLCIQRWNWGPAFPPRPGLAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSSIGSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529  1879  A 4039  2 366  KDMVLIMEMQSMITMKCPQYL*E*	LIKDYMLEN	DVAVYFIIKEWAIMGPAEKALYKDYN			1			
528 1878 A 4028 1160 242 GTSELLCIQRWNWGPAFPPRPGLAI VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSS\GSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	QUKESCISPA	YGGCGPL*CHPTSKPALVFS\LEQUKES				1	1	
VEMGSAKSVPVTPARPPPHNKHLA PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSSIGSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	RYTYLS	TGSSLSRNDWRAGWIGYLELRRYTYLS		l			_	1
PSAGILRTPIQVESSPQPGLPAGEQL DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSSIGSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*			242	1160	4028	A	1878	528
DSDPRSPTLGIARTPMKTSSGDPPSI VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSSIGSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*						1	-	Í
VFETEDSKSNLPPEPVLPPEAPLSSE QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDFETPRSS\GSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LL WEQGQD\HDKENQHFPLVES 529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	EQLEGLKHAC	PSAGILRTPIQVESSPQPGLPAGEQLEGL	ļ			1	ĺ	
QLSVEEQMPPWNQTEFPSKQVFSK ETPVASQSSDKPSRDPETPRSS\GSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES  529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	PPSPLVKQLSE	DSDPRSPTLGIARTPMKTSSGDPPSPLVI					-	1
ETPVASQSSDKPSRDPETPRSS\GSM NSSKVL\GKSPLHPSCQDDNSPGTL AFKPLSENVSELK\EGA\ILGTGR\LI WEQGQD\HDKENQHFPLVES 529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*							1	
NSSKVL\GKSPLHPSCQDDNSPGTL	FSKEEARQPT	QLSVEEQMPPWNQTEFPSKQVFSKEEA		ļ		1	i	
AFKPLSENVSELK\EGA\ILGTGR\LI   WEQGQD\HDKENQHFPLVES   529   1879   A   4039   2   366   KDMVLIMEMQSMITMKCPQYL*E*				İ	ļ	1		1
WEQGQD\HDKENQHFPLVES   529   1879   A   4039   2   366   KDMVLIMEMQSMITMKCPQYL*E*	GTLTLRQGKA	NSSKVL\GKSPLHPSCQDDNSPGTLTLR	ł	1		1	ŀ	1
529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*	R\LLKTEGRA	AFKPLSENVSELK\EGA\ILGTGR\LLKTE				1		
				1		1		
CW*GCGSTGILIFC/WS*PL*KTI*OI	L*E*RKIPDITK	KDMVLIMEMQSMITMKCPQYL*E*RKI	366	2	4039	A	1879	529
, , , , , , , , , , , , , , , , , , ,	I*QPR*FKQI,*T	CW*GCGSTGILIFC/WS*PL*KTI*QPR*F	1		1	1		
		ILTIIYSIM*EHTFHNAGV*LSDIYPRFMI	1	}		ļ		1
HTEICT*MFIAVLFVVVKTWKQF				1				1
		LLEVNGNTIVTVFTKAQNKKNKGSRSI	1 3	358	4057	A	1880	530
		RKYGSRINLLKSKHDKNICTENYKT*M	1	1 333	1007	^	1000	1 220
		/DTDKWKDILCSWIRRIHMKDILCSWIG	1	1			}	1
VKISILPKVNYRFYLISIKIIMAI				1				1
TO OTTORING A SPECTAL M		TOGTEEIYKISSCEWVQASFSTPLITLHI	278	150	4061	+_	1001	521
		HKATVIKMVWYWHRQ*KFSKN/RIESS	2.10	1 30	4001	^	1991	331
		IYDOFIFDKGEKIIQEKGNSFFNN/MCW	1		1		l	1
T*KR					1	1		1
A STATE OF THE PROPERTY OF THE	TVTEVETOO	NDLLENFKFWE*FKE*LENINGTVTEKI	1260	10	1000	+	<del></del>	
		YKELSSPKYSGTRQFYGQTISNFPGKIIS	308	119	4069	A	1882	532
			1	1				
	TI I I LINK DINI YI	KLFQNTE/TEGRHPISLYEFRITLITIPNK	}			1	1	1
QIWMPVSLMNIVTLKCPT	TIND DOOR TO							L_
1 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	· ·	PIRKFTKVAG*KSNTPK*LAFLHINNEQ	355	1	4076	A	1883	533
		ITNI/PFIIASKRIKYSGISLTKEMKDLYT		1				
KIKEDI'NKWKDI/SCFWVGR/LNIV	'VIAKWLK\AIG	KIKEDTNKWKDI/SCFWVGR/LNIVKMI			_1_			1

CEO ID	CCO ID	Mat	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID NO: of	Met hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of		поц	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	09/496		to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	Ì		correspondi ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		[	914			T=Threonine, V=Valine, W=Tryptophan,
		1		amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
l	[	ļ	1.	residue of	sequence	
		]	1	peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion IFNAIPIKMPMMCMAKIEKNSS
-25-1		<u> </u>	1000			
534	1884	A	4088	3	1931	IIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC
	Į					GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL
	[	Į	1 .			QTRLVDAAKALNLVHCHCLDIFINQAFDMQR
	1	1				DLQITPKRLEYTRKKENELYESLMNIANRKQE
	ĺ	l	ì			EMKDMIVETLNTMKEELLDDATNMEFKDVI
	1	1	1	4		VPENGEPVGTREIKCCIRQIQELIISRLNQAVA
	)	]	1		]	NKLISSVDYLRESFVGTLERCLQSLEKSQDVS
	İ	l	İ		İ	VHITSNYLKQILNAAYHVEVTFHSGSSVTRM
		l	ľ			LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI
						ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ
		ŀ			1,	LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS
i	1	1	1	ł	1	LESRSLQDVLLHRKPKLGQELGRGQYGVVYL
	1	ļ		1		CDNWGGHFPCALKSVVPPDEKHWNDLALEF
		1				HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA
	ļ	ļ		}	ļ	VLLIMERLHRDLYTGLKAGLTLETRLQIALDV
,						VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI
•			l			TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK
				1	,	YDNSVDVYAFGILFWYICSGSVKLPEAFERCA
			ļ	ļ		SKDHLWNNVRRGARPERLPVFDEECWQLME
		1	1 .	ľ		ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\
				ŀ		NSEQPNRGLDDST
535	1885	A	4090	2	417	ALMPHEANYEEIFLKTDKDMDGFESGLEVRE
				1		IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD
	1		1	1	• .	HFALAFHLIT\QKLIKGIDPPLVLTPEKISPSNR
	1	ĺ	1	į.		ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK
ı			İ	ļ	ł	HNRKRIWLRA
536	1886	A	4102	569	829	DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI
						PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK
į	1	1		}		EQNLEESHYLDFK*YYRAV
537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML
55,	1007	** .	1	1	1	IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM
	1 .	]	J	1		HSLGIDQQKTIE
538	1888	A	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL
336	1000	^	7105	***	314	VPRENITAWMNAIGLIITALPVS
539	1889	A	4111	268	<del>                                     </del>	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM
337	1007	1^	7111	200	1	I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW
	1		}	1	1.	PLFPKIYREGTLPHSFYEASITL
540	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE
J4U	1920	I A	4142	170	2004	
	1	1	1	1	1	VRGHGGVRESGRAPQQPGRRRGRRPRKRPR
		i			]	GRWRREGCGAGGRGVCVAAWSORSIAGNN
1	1	1	}	1	1	
	1			1		DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH
		1		1	1	IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR
· .	1	}			1	CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT
1			1	1	1	MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY
İ	}	1				GFPELEDSTSEYLCTILNIQNVCMTFDVASLY
		1	1		1	SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK
}	1				1	TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK
			ľ			
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF
						ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HAESENFAFWQDMKWKNKFWGKSLEIVPVG
						TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG PTPGGOCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR NHLIFRGGAQIIFLATFDDSPKAVLGDRLLLT ANVSSENNTPRTSKTTFQLELSVKDAVYTVV SSH
543	1893	A	4153	678	11	TISYPQCLTQMYFLISFANVDTFLLPIMALDH YVAICSALQ*CSIITP/ELCQGLPVLA*AGSSLIS PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA AAILRIPSPTRRKACSICSSHLSLVTLFYGTV LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3		LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF
545	1895	A	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA GITGVSHHA*PRASGRRCW
547	1897	A	4176	3029		AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASVPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPAISPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGPTSP MP*CFHRPSPPLP/LSSPFPA/LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		1	1	amino acid	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	<del> </del>	<del> </del>	<del> </del>	soquenee		GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA
	İ		1		[	PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ
				ţ	İ	VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP
			1			GPPGLPPLQSSLSFPPPPPPPVPQPPAPPALQWG
		i			1	LHLPGGRTK
548	1898	A	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK
						KIQFHQELLVLFWKLCDFNKVGQPRGALQGD
		}		Į.		GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG
						PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR
		ļ		l	1	ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP
	i					YSIRVPMDIPVFTGTHADLLIV\VFHKIITSGHQ
						RLQPLFDCLLTTVVNVSPYLKSLSMVTANKLL
						HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI
	ŀ	1	į			IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP
	Į	}	į	İ		TIHKALQRRRRTPEPLSRTGSQGGAPPWRAPA
	[		1		i	PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG
		]	1			SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ
		) :	ŀ	j .	ļ	TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ
	1	1			İ	HGTLVGLLPVPHPILIRKYQANSGTAMWFRT
	1	}		ŀ	ŀ	YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT
J43	1077	^	71/1	550	321	ALLODGRRKVHYLFPDGKEMAEEYDEKTSE
		1		1	ł	LLVRKWRVKSALGAMGQWQLEVGDPAPLG
		1		1	İ	AGNLGPELIKESNANPIFMRKDTKMSFQWRIR
			1			NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK
					1	KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	A	4192	1	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL
						GASAMRRSEVLAEESIVCLQKALNHLREIWE
						LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE
		i				SLKERLIKSISVCQKELNTLCSELHVEPFQEEG
			i	1	1	ETTILQLEKDLRTQVELMRKQKKERKQE\LKL
						LQEQDQELC\EILCMPHYDIDSASVPSLEELNQ
				1		FRQHVTTLRETKASRREEF/VSSIKRQIILCME
	1:					ELDHTPDTSFERDVVCEDEDAFCLSLENIATL
		1	]	J	ļ	QKLLRQ\LEMQKSQNEAVCEG\LRTQI\RELW
	ľ	1	1	ļ		DRLQIPEEEREAVATIMSGSKAKVRK\ALQ\LE
						VDRLEELEKCKTMKKVIEAIRVELVQYWDQC
						FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER
		1				KASDPNRFTNRGGNLLKEEKQRAKLQKMLP
		1				KASDPARFTARGGALLREER QRANLORMLP KLEEELKARIEL WEQEHSKAFMVNGQKFME
	1	1				YVAEQWEMHRLEKERAKQERQLKNKKQTET
	1			1	İ	EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT
	i	1				TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK
}		1	1			PVAASTCSGKKTPRTGRHGANKENLELNGSI
			1	į		LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS
		1	1			DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	A	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV
1 33.	1,01	1.,	1	1	1	VALLDICFILVAVPESLPEKMRPVSWGAQISW
1						KOADPFASLKKVGKDSTVLL\ICITVCLSYLPE
		1	1	l		AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI
l				1	l	LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML
		1		J	1	QLAWYGFGSQAWMMWAAGTVAAMSSITFP
	1	-	1		1	AISALVSRNAESDQQGVAQGIITGIRGLCNGL
	1	- [	1	1		GPALYGFIFYMFHVELTELGPKLNSNNVPLQ
1	1	1	1	1	1	GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS
ļ			l l	Į.	į.	QAVIFOFFE GACIVENIOF EVALUE FOR AS
	ļ.			1		GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS

PCT/US01/03800

SEQ ID   SEQ ID   Met   SEQ   Predicted   Predicted end   Amino acid sequence	ce (A=Alanine C=Cysteine,
NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E	
nucl- peptide in nucleotide location F=Phenylalanine, G	=Glycine, H=Histidine,
cotide seq- USSN location corresponding I=Isoleucine, K=Ly	
	Asparagine, P=Proline,
uence 914 ng to first acid residue Q=Glutamine, R=A	rginine, S=Serine,
	aline, W=Tryptophan,
	known, *=Stop codon,
	le deletion, \=possible
sequence nucleotide insertion	
552 1902 A 4197 2 14302 ARPPPAPGSRQQI	KQKAAPGAAAAELRGAR
	DGGEGEDEIQFLRTDDEVV
	KLCLAAEGFGNRLCFLESTS
	TVLEQSLSVRALQEMLANT
	WKFMMKTAQGGGHRTLL
, I I I I I I I I I I I I I I I I I I I	GMYLCCLSTSRSSTDKLAFD
	WWTIHPASKQRSEGEKVR
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ERYLHLSYGNGSLHVDAAF
f	GSEAAQGYLIGGDVLRLLH
	GEHGEEQRRTVHYEGGAVS
	RVAWSGSHIRWGQPFRLR
	EDKNLLLMDKEKADVKSTA
	GVRKEVDGMGTSEIKYGDS
	VLTYQSVDVKSVRMGSIQR
	DGISLSRSQHEESRTARVIRS
1 1 1 1 1 1	DALSKKAKASTVDLPIESVSL
	DEHLEHEDKONRLRALKNR
	VLECIDRLHVYSSAAHFAD
	SILNSLYELLAALIRGNRKN
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	SRLERLEASSGILEVLHCVL
	HIKSIISLLDKHGRNHKVLD
	AVRSNQHLICDNLLPGRDLL
LQTRLVNHVSSN	/RPNIFLGVSEGSAQYKKWY
YELMVDHTEPFV	VTAEATHLRVGWASTEGYSP
YPGGGEEWGGN	GVGDDLFSYGFDGLHLWSG
CIARTVSSPNQH	LLRTDDVISCCLDLSAPSISF
RINGQPVQGMFF	ENFNIDGLFFPVVSFSAGIKV
RFLLGGRHGEFK	CFLPPPGYAPCYEAVLPKEKL
· KVEHSREYKQEF	RTYTRDLLGPTVSLTQAAFT
PIPVDTSQIVLPP	HLERIREKLAENIHELWVMN
KIELGWQYGPVI	RDDNKRQHPCLVEFSKLPEQ
ERNYNLQMSLE	TLKTLLALGCHVGISDEHAE
DKVKKMKLPKN	IYQLTSGYKPAPMDLSFIKLT
	ENAHNVWARDRIRQGWTY
	PRLVPYTPLDDRTKKSNKDS
	GYNLEAPDQDHAARAEVCS
GTGERFRIFRAEI	KTYAVKAGRWYFEFETVTA
	GCQPDQELGSDERAFAFDGF
	HYGRSWQAGDVVGCMVDM
	EILLDDSGSELAFKDFDVGD
	VGRMNFGKDVSTLKYFTIC
	NTNRDITMWLSKRLPQFLQV
	OGTIDSSPCLKVTQKSFGSQN
	PIECAEVFSKTVAGGLPGAG
	DADSDFEVLMKTAHGHLVP
1 1 1 1 1 1	PEFNNHKDYAQEKPSRLKQ
	STSHSARLTEDVLADDRDDY
	SVRIFPGQEPANVWVGWITS
	DRVRTVTVTLGDEKGKVHE
	AGESMSPGQGRNNNGLEIGC
	IANGKELSTYYQVEPSTKLFP
	FQFELGRIKNVMPLSAGLFKS
	RLHVQFLSHVLWSRMPNQFL
	WLVQCLDPLQFMSLHIPEEN
	ELLKFHYHTLRLYSAVCALG
	IVDEPQLLYAIENKYMPGLLR
	SSYATARLMMNNEYIVPMT
	NKKHGLPGIGLSTSLRPRMQF
	QYSPEFPLDILKSKTIQMLTE
AVKEGSLHARD	PVGGTTEFLFVPLIKLFYTLLI

PCT/US01/03800

- CEA TA	OEA ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID				nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning		F=Phenylalanine, G=Glycine, H=Histidine,
nucl~	peptide		in	nucleotide	location	r-rachylmanme, G-Glychie, ri-rashume,
eotide	seq-	<b>.</b>	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
l	l .	1	i	peptide		/=possible nucleotide deletion, \=possible
{	ĺ		1	sequence	1	nucleotide insertion
<u> </u>				Sequence		MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
•	1	1	ì		ļ	LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL
1	1	ł	ļ	İ		
	ļ	Į.	ĺ	Ì	1	QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
1	ł	ł	ł	i		VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
l	1	1 .	l .		ļ	ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
i			1		1	CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG
1	1	1	1	ì	l .	NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
i		1	Ì	ļ		DSKKSSTLQQLISETMVRWAQESVIEDPELVR
İ	1	]	j	J	<u> </u>	AMFVLLHROYDGIGGLVRALPKTYTINGVSV
1	1		1		1	EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
ŀ	İ	1	ł			LGDIMNNKVFYQHPNLMRALGMHETVMEV
1			l .	ļ		MVNVLGGGESKEITFPKMVANCCRFLCYFCR
1	1		1	1		
	1	1			1	ISRQNQKAMFDHLSYLLENSSVGLASPAMRG
ł	1		1	1	•	STPLDVAAASVMDNNELALALREPDLEKVVR
1	}	1	١.			YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
1	1	İ	1			YLDFLRFAVFCNGESVEENANVVVRLLIRRPE
				-	•	CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
1	1					GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY
-	.				ŀ	SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
1	İ		1	1	1	LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
1	ł	1	1	ł	1	AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL
			1	1		EVGFLPDLRAAASLDTAALSATDMALALNRY
1	ŀ	1	1	1		LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT
	•	1		1	1	VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
1	1	1 .	1	Ì	ì	
1		·	J	1	]	MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
1	1	1	1	1	1	YERCWKYYCLPGGWGNFGAASEEELHLSRK
1		ļ	1		İ	LFWGIFDALSQKKYEQELFKLALPCLSAVAG
		1	1		1	ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ
1	1					PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK
1			1	1		LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE
1	ł	1	i	1	1	KEIYRWPIKESLKTMLARTMRTERTREGDSM
	ļ			1		ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
1	1				1	NVTLSRDLHAMAEMMAENYHNIWAKKKKM
1		1	1	1		ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
			1	Ì	į	QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
1	1	}	1	1		COLLEGING IN A SKOLK DECEDIT SECRETA
}	1	1				YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF
1		1	1		1	PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA
_		1		1		SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
1		1		1		RISLFGNDATSIVNCLHILGQTLDARTVMKTG
				1		LESVKSALRAFLDNAAEDLEKTMENLKQGQF
1	1		1	1	1	THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
	1	1	1	}	1	GQHQFGEDLILEDVQVSCYRILTSLYALGTSK
	1				1	SIYVERORSALGECLAAFAGAFPVAFLETHLD
	1		1	1		KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
1	1	1	1			SLEKLMEEIVELAESGIRYTQMPHVMEVILPM
1	1		1			LCSYMSRWWEHGPENNPERAEMCCTALNSE
1	1		1	1	1	HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF
1	1	ì	1	1	1	
1		i	1			SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
			1	1	-	SEEDHLKAEARGDMSEAELLILDEFTTLARDL
	i	ļ	1	Į.		YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
			į		1	MVAEVFIYWSKSHNFKREEQNFVVQNEINN
						MSFLITDTKSKMSKAAVSDQERKKMKRKGD
		1	1		1	RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
	1		1			LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP
ł	ļ	1		1		AIRWQMALYKDLPNRTDDTSDPEKTVERVL
	1	1	1			DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK
	1			1		KAVWHKLLSKORKRAVVACFRMAPLYNLPR
1	1	1	{	i	1	
				}		HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
		1	1	1	1	KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT
			1	<u> </u>		EKCKLEEDFLYMAYADIMAKSCHDEEDDDG

NO. of No. of Poptide ox			N 6-4	Long	Dendicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
in undectide sequence of the corresponding plant	SEQ ID	SEQ ID	Met	SEQ	Predicted		D=A grantic Acid F=Glutamic Acid
Sequence   1914   1914   1915   191			noa				E-Phenylolonine G-Glucine H-Histidine
uence unce by the corresponding the properties of the control of t		• •	1				
### Big of first amino and residue of peptide sequence peptide peptide peptide peptide peptide peptide peptide peptide peptid	eotide	seq-	ļ				
mino seld residue of peptide sequence peptide sequence of peptide sequence	seq-	uence	1				
residuo of peptido sequence	uence	1	1	914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
Popsible mueleotide diestrion   Popsible mueleotide insertion   mueleotide insertion   mueleotide insertion   mueleotide insertion   mueleotide insertion   mueleotide insertion   mueleotide insertion   mueleotide insertion   mueleotide   mueleotide insertion   mueleotide   mu	[	1	1	1	amino acid		T=Threonine, V=Valine, W=Tryptophan,
Sequence   mucleotide insertion	<b>(</b>		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Inucleotide insertion	Ì		1	Ì	pentide	l -	/=possible nucleotide deletion, \=possible
EEVKSFEEKMEKOKILIYOQARILIDROAA BMYLOTISASKGETIPOWAATIKIGIAILNOG NSTVQOKMLDYLEKKEDVGFFQSLAGLMOS CSVLDIAAFERONKAGIGAMTEGSGEKV LQDDEFTCDI.PET.QLLCEGINSDFQNYLRT QTONNTTVNIISTVDYLEVGESISDFWYYY SGKDVDEQOQRNFSKAIQVAKQVPINTLEYY QGETGNQGSLASISLWDAVVGFILHVFAHM QMKLSQDSSQIELIKELMDLQKDMVVMLIS MLEGONVONTIGKOMYDMI VESSNIVEMIL KFPDMFLKLKDLTSSDTFKEVDPDGKGVISK ROPHK-AMESIKHYTYGSETELISACATIDENE TLDYEEFVKRYHEPAKDIGFNVAVLITNLSHI MPNDTILOTIFLIASSVINTYOPHTEMQLAA QUSESDLNBRSANKESEKERPEGQPBMAFF SILTYMSALFALRYNTHMILTIMBMISLIKSIKKOM KVKKMITYKDMYTAFTSSYWSIFMTLLHFV ASVERGEREKKERJENGTBERGLGVBLAF SILTYMSALFALRYNTHMILTIMBMISLIKSIKKOM KKVKKMITYKDMYTAFTSSYWSIFMTLLHFV ASVERGEREKKERJENGTBERGLGVBLAF SILTYMSALFALRYNTHMILTIMBMISLIKSIKKOM KKVKKMITYKDMYTAFTSSYWSIFMTLLHFV ASVERGEREKERJENGTBERGLGVBLAF SILTYMSALFALRYNTHMINTLANGTBERGYKLIF ANMEDPTQDEVKDDGEEGERFPLEALPSBAL ANMEDPTQDEVKDDGEEGERFPLEALPSBAL KKVKKMITYKDMYTAFTSSYWSIFMTLLHFV ASVERGERFENGTENGTBELLGSSUFGAKKIKVAELL ANMEDPTQDEVKDDGEEGERFPLEALPSBAL STENDER SILTYMSALFALRYNTHMINTLANGTBERGYKLIF ANNENDETTENDELLGSSUFGAKKIKVAELL ANMEDPTQDEVKDDGEEGERFPLEALPSBALFTLHITYDS FRINTATION FOR THE TOP TO THE TO	i	į	1	1	,	[	
EMYLOTISASKGETGMVAATLKLGIALINGG NSTVQCKMLDYLKEKKDVGFFGLAGM\QS CSVLDIAAFERQNKAEGLGMYTEGGGEKV LQDDEFTCDLFRFLQLLCGGINSDFONYLRT QTGNNTTVNIISTVDYLLRVQBSISDFYWYY SCKDVDDEQQQRNFSKAIQVAKQVPNTLTEYT QGPCTGNQQSLABISLWDAVVGFLHVFAHM QMALSQDSSQIBLLKELMDLQKDMVMLIS MLEGNVVNGTGKQMYDMLVESSNNVEMIL KPFDMFLKLDLTSDTFKEVDPDGKVISK RDFHKAMESHKHYTQSETEFLJSCAETDENE TLDYSEFVKRPHERAKDIGFNVALVLINLISH MPMDTRLOTTELASSVLNYOPFLGREIMG SAKRIERVYTEISESSTTQWKEVQVESKROFI FDVVNEGGEKEKMELVNYCFDTIFEMQLAA QUSSIDLINERSANGESESKERPEQOPPMAFF SILTVRSALFALRYNILTLAMMLSLSSLKKOM KKVKKMTVXDMYTAFTSYWSIFMTLLIFFY ASVFRGFFRICSLLIGGSULFGGAKKVAELL ANNEDPTQDEVRGDGEGEGRFLLEAALPSED LTDLKETFESDLISJDFGLDLRRGGQYKLIP HNPNAGLSDLMSNYVPMEVQEKFQEQYKLIP HNPNAGLSDLMSNYVPMEVQEKFQEQXK EEKKEKETKSEFEKKAEGDGGEKEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYNMMALELVAFAINELL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHBIJAVHYVLESSGVMEPTVRILPLILHIVISF FCIIGYYCLKVPLVFKREKEVARKLEFDGLTY TEOFEDDIKGQWBULVNTQSFPNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKFKKDKSSLSAVINSIDVXQMW KLGVVFIDNSRLVLAWYMT  1903 A 4199 31 767 LPENGGGGGCGCAERTQGAARGSAAA GGLYLRDKGRUCCINSBEGGGGRETQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTGVAYA LPLSKGRGAGTRRAFFSRGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYLRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGAERTQVAA GGLYRDKGRUCCINSBEGGGGAERTQVAA GGLYRDKGRUCCINSBEGGGGAERTQVAA GGLYRDKGRUCCINSBEGGGGAERTQVAA GGLYRDKGRUCCINSBEGGGGGAERTQVAA GGLYRDKGRUCCINSBEGGGGAERTQVAA GGLYRDKGRUCCINSBEGGGGGAERTQV		<del> </del>	<del> </del>	<del> </del>	Soquence		FEEVKSFEEKEMEKOKLLYOOARLHDRGAA
NSTYQQKMLDYLKEKKDVGFFQSLAGLMQS CSVLDLAFREQNKAGLGAWTEGSGEKV LQDDEFTCDLFFEQLLCEGINSDFQNYLRT QTGNNTTVNIISTVDYLRVQESISDFWYY SGKDVUDEQQQRNFSKAIQVAKQVFNTLTEYY QFCTGNQGSLASISLWDAVVGFLHFVAHM QMKLSQDSSQIELLKELMDLQKDMVVMLIS MLBGNVVNOTIGGUMYDMLVSSNNVEMIL KFPDMFLKLKULTSSDTFKEVPDPOGKOVISK RDFHKAMESIKHYTQSETEFLJSCAETDENE TLDYBEVKRFHEFAKDIGFNVAVLTINLSEH MYPDTRLQTFLELASSVLNYFQFFLGRIEMG SAKRIERVYFEIGSSKTQWKFQVYLGSKRQFI FDVVNEGGEKKMELFVNFCEDTIFEMQLAA QISESDLNERSANLESSKTQWKFQVYLGSKRQFI FDVVNEGGEKKMELFVNFCEDTIFEMQLAA QISESDLNERSANLESSEKEPEPEGGFRMAFF SILTVRSALFALRYNILTLIMKMISLKSLKKQM KKVKKMYVZDMYTAFFSSYWSIFMTLLHFV ASVFROFFRICSLLLGGSLVGGAKKKVAELL ANMPDTTQDEVRGDEGEKRFLEADEGGYKLIP HNPNAGLDLMSNPYPMFPVQEFFQGXAK EEKEKEETKSEPEKAEGGOTEKERAKED LTDLKELTESDLLSDIFGLDLKREGGGYKLIP HNPNAGLDLMSNPYPMFPVQEFFQGXAK EEKEKEETKSEPEKAEGGOTEKERAKED LTDLKELTESDLLSDIFGLDLKREGGGYKLIP FYKVSTSSVVEGKELFTRSSSPNAKVSLDS SHRILAHFYVLESSGYMEPTVRILLHTVISF FCIIGYYCLKVPLVFKEKEVARKUSLDSS SHRILAHFYVLESSGYMEPTVRILLHTVISF FCIIGYYCLKVPLVFKEKEVARKUFFLDLYT TEOPEEDDISGGWDRLLYDTGSFPNNYDDFF VKRKVMDKYGEFYCRDRSSELGMAKKAALL FSTANDERSKEKFKCFKDSSLSAVLNSIDVKYQMW KLLGVVFINNSRLLAWYMI  553 1903 A 4199 31 767 1767 1767 1767 1767 1767 1767 17	1			ļ.			EMVI OTISASKGETGPMVAATI KI GIAILNGG
CSVLDINAFERQNKABGLIGMYTEGGGGEK LQDDETCDLPRIQLICGEGINBSTOPNYLRY QTGNNTTVAIIISTVDYLLRVQBSISDFYWLRY QTGNNTTVAIIISTVDYLLRVQBSISDFYWLRY QTGNNTTVAIIISTVDYLLRVQBSISDFYWLRY SIGKDVDERQGRYRSKAIQVAKQYFNILTEYI QGPCTGNQQSLAHSRLWDAVVGFILIVAH QMKLSQDSSQIELKELMDLQKDMYVMLLS MLEGNYVNGTIGKQMYDMLVESSNNVEML KPFDMPLKLKDLTSSDTFKEVDPDGKGVISK RDFHKAAMESHKHYTOSETEFLLSCAETDEM TLDYEFPVREHPEARDIGFNAVALITNI.SSH MPNDTRLCIFLELAESVLJYFQPFLGRIEMG SAKRIEVYFEISESSRTQWHERDYESKRQFI FDVVNBEGGEKEKMELFVNFCEDTTEBMLQAA QISEDDLNBRANKEESEKERPEQGRMAFA SILTVRSALFALFVHILTMRMLSLKSLKKQM KKVKKMTVKDMVTAFFSSYNSIFMTLLEFY ASVFRGFRIICSLLGGSLVBGAKKKVAEIL ANMPDFTQDEVRODGEGEGKVLLAALPSD LTDLKBLTESEDLLSDSFLDLKRRGGQYKLLP HNNNAGLSDLMSNPVPNBEVQEKFLGGXAK EEKKEKEETSEFKAGGGDGKKEKAAED KGKQKLRQLHTHRYGEFEVESSFAWKVISDDS SHRIIAVHYVLEESFGKAGGGDGKEKAKAED KGKQKLRQHTHRYGEFEVESSFAWKVISDDS SHRIIAVHYVLEESGGYMEPTVRUPLHITVISS FCIIGYYCLKVPLVFRREKEVARLEFDGLTY TEQPSEDDIKQOWRDLVNTOFSHVUDKT VKRRVMDKYGFFYGRDRISELLGMDKAAH FYKNSTSSVVBGKELPTRSSSENAKVISDDS SHRIIAVHYVLESSGGYMEPTVRUPLHITVISS FCIIGYYCLKVPLVFRREKEVARLEFDGLTY TEQPSEDDIKQOWRDLVNTOFSHVUDKT VKRRVMDKYGFFYGRDRISELLGMDKAAH FYKNSTSSVVBGKELPTRSSSENAKVISDDS SHRIIAVHYVLESSGGYMEPTVWBFT VKRRVMDKYGFFYGRDRISELLGMDKAAH LPLSHGHSIGGGGCRCAAEN/GAARGSAK GUIMGEDDDSHESMRLYKNIDTVSQMW KLGVVFTDNSFLYLAWYM LPLSHGHSIGGGGCCAAEN/GAARGSAK GUIMGEDDDSHESMRLYKNIDTWSTQMV REPOTTSLNRDLTGGLEYATKISRSFSNYFGW AYGLYLBIDKGLQCLNESSEGGGRGVFKFW BRADDRSKFYESDADBELLFNIPTTGHVAVA AYGLYLBIDKGLQCLNESSEGGGRGVFKFW BRADDRSKFYESDADBELLFNIPTTGHVAVAP RSSQCKNTTGVYFYGLGRGBVTLAYQAP RSSQCKNTTGVYFYGLGRGBVTLAYQAP RSSQCKNTTGVYFYGGGGCCCAAEN/GAARGSAK HEISDATA UTTRAGARAPADHRYHQVTPQTHTHS  554  1904  A 4200  1 961  GIPCTEMGRGBANDANVTGGELEFATKISRSFSNYFGG TAQLRGSSLELTVWDQALFGMADRILGD RVRLARGALTAKALVLPSRTRALGACGGOTDQYSL HQULCLVVLGAKRIFVYGMSFDTLX UTTRAGARADARSTGNGGVYGKSSEDGGSETLPQDH NHEMERRERWQQERLHREAYYQFINELBHD VKLMDHALLGTGGITSELQQRLDGVKE QLASQPDLRDGTMYFDSBVPRESSHEDSLLB  WTDMTLVLH  BENKKARNLGGTGAACSGKCWGWLSSFN LWTDMTLVLH  CHRKKARNLGGTGAACSGKCWGWLSSFN LWTDMTLVH NNGEFRSLEHIVNIERHORGEHGEDYTUDL DNRHAMBORGRSTGFVAR	}	i	1	1	<b>!</b>	ŀ	PICTUOUNI DVI KEKKDYGEEOSI AGI MOS
LQDDEFTCDLFRFLQLLCGGHNSDFONYLY GTGNNTTV-NIBITYDVLLKVQSSISDFYWYY SGKDVDBGQGRNFSKAIQVAKQVFNITERY QGCTGNQGSLABRILWADVGGLFUFAHM QMKLSQDSSQIELLKELMDLQKDMVVNLIJ MLEGNVVNGTIGKQMVDMLVSSSNNVEML KFPDMFLKLKDLTSSDTFKEYDPDGKGVISL MLEGNVVNGTIGKQMVDMLVSSSNNVEML KFPDMFLKLKDLTSSDTFKEYDPDGKGVISL RDFHKAMESHKHYTOSETEFLSCAETDENE TILDYEEPYVREPHEPAKDIGFNVAVLLTNIS. RDFHKAMESHKHYTOSETEFLSCAETDENE TILDYEEPYVREPHEPAKDIGFNVAVLLTNISL SAFRUFFTIGSSRTQWEKPQVKESKRQFI FDVVNBGGBKERMELFVNFCEDTIERMGA GARRENYFEISESSRTQWEKPQVKESKRQFI FDVVNBGGBKERMELFVNFCEDTIERMGA KKVKKMTVKDMVTAFFSSYNSFMTLLHFV ASVFRGFFRIGSLLLGGSLVEGAKKIKVALL ANMPDPTQDEVRGDGEGEGREKAKED LTDLKBLTESSGLVBGAKKIKVALL ANMPDPTQDEVRGDGEGGERKAKED LTDLKBLTESSGLVBGAKKIKVALL HNPAGLSDLMSNPVPMPEVQEKFQEGKAK EEKEEKEEKETSSPERABGEDGEKEEKAKED KGKQKLRQLHTHRYGFEPVPSKFKKILDY QQKLLNYFARNFYNNRMLALFVAFANNFIL FYKVSTSSVVEGKELPTRSSSEAFKKILDY QQKLLNYFARNFYNNRMLALFVAFANNFIL FYKVSTSSVVEGKELPTRSSSEAFKKILDY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVVEGKELPTRSSSEAFKKILDY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVVEGKELPTRSSSEAFKKILDFOLLAY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVEGKELPTRSSSEAFKKILDFOLLAY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVEGKELPTRSSSEAFKKILDFOLLAY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVEGKELPTRSSSEAFKKILDFOLLAY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVSTSVEGKELPTRSSSEAFKKILDFOLLAY VARRKVMDKYGFFVGRNDSLLGMDKAALD FYXTSTSVSTSVEGKELPTRSSSEAFKGGAERTQVAA AYGLYLRIDKGRJCCALBESSGGGVFKFW ERADDRSKFYESDADELLFNPFTGGHVLA GIIMGEDDDSHFSMRLYKNIDMSFDTER EPDQTFSLNRDLTGLELFYATKISRSSNYHF BENKGRADTTKYFYIGLRGEWFLERREHEVTI CNYBASAPADHRVHQVTTQTHFIS HISKNGADTTKYFYIGLRGEWFLERREHEVTI CNYBASAPADHRVHQVTTQTHFIS HISKNGADTTKYFYIGLRGEWFLERREHEVTI CNYBASAPADHRVHQVTTQTHFIS HISKNGADTTKYFYIGLRGEWFLERREHEVTI CNYBASAPADHRVHQVTTQTHFIS HISKNGADTTKYFYIGLRGEWFLERREHEVTI CNYBASAPADHRVHQVTTQTHRIS HISKNGADTKYFYIGLRGEWFLERREHEVTI CNYBASAPADHRVHGUTTQTHRIS HISKNGADTKARLPYPGERCHEDFTTGH WTDMTLIVH BENKKARNILGHGGERLAYPGFRISHPSTGCH TPAQLRQSSLELTVWDQALFGGMTGNFFKGCL LPDQQKLRLEREAYYQFINELND DVRLMDDHALLGTFGGETIERDDVTGH NHEMERER WQQBRLHREBAYYQFINELND DVRLMDDHALLGTFGGETIERDD	1	1		1			NSI VQQXVILD I LKEKKD VOI TQGLACLINQS
OTGNNTTVMIISTVDYLLRVQESISDFYWY SIGKDVDEQGGRNFSKAIQWAKQYFNILTEYI QGPCTGNQQSLAHSRLWDAVVGFILIVPAM QMKLSQDSSQIELL KELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVESSNNVEMI KFPDMFLKLKDLTSSDTFKEVPDGKGVISK RDFHKAMESHKHYTQSETEFLLSCAETDENE TILDYEEFVRRHEPAKDIGRWAVLITINESSNEM TILDYEEFVRRHEPAKDIGRWAVLITINESSNIPKEND SARRIERVYFEISESSRTQWERCPQVKESRK, OR SARRIERVYFEISESSRTQWERCPQVKESRK, OR SARRIERVYFEISESSRTQWERCPQVKESRK, OR SARRIERVYFEISESSRTQWERCPQVKESRK, OR SARRIERVYFEISESSRTQWERCPQVKESRK, OR SARRIERVYFINLTIMRMISLKSLKKQM KKVKKMTVKDMVTAFFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGDGEGERKPLEAALPSS LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGDGEGERKPLEAALPSS LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGDGEGERKPLEAALPSS LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGDGEGERKPLEAALPSS LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGDGEGERKPLEAALPSS LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGDGEGERKPLEAALPSS LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGGEGERKAED LITDKESTEFSSYWSIFMTILLFY ASVFRGFRIEGLILGGSLVGGAKKKVAELL ANMEDPTODEVRGGEKKKKTPLEAALPST LITTKTSSSENAKVISLOSS SHRIIAHTHYVLESSGYWEFTVERLIPLHTVIT LITTKTSSSENAKVISLOSS SHRIIAHTHYVLESSGYWEFTVERLIPLHTVIT FORSCOLKYTTONSTYLAWFMT TEQPSEDDIKGQWORLVINTOSFFNNYVONY KLGVYFUNSEYLVAWFMT TEQPSEDDIKGQWORLVINTOSFFNNYVONY WKLGVYFUNSEYLVAWFMT LITTKTSSSENAKVISLOSSGRGFRYFXW BRADDRSKFYRSSOLDEELLIPTGHVKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGEDDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGERDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGERDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGERDDSHESMRLYKNIDQMSFDOTFGHWKLK GIIMGERDDSHESMRLYKNI	1		1	1			CSVLDLNAFERQNRAEGLGWY I EEGSGERY
SGRDVIDEOGGRNFSKAIQVAKQVFNITLENG QGPCTONQGSLABRIL MADVGELHVRAHM QMKLSQDSSQIELLKELMDLQKDMVVMLIS MLEGNVVNGTIGKQMVDMLVSSSNNVEML KPFDMFLKLKDLTSSDTFKEYDPDGKGVISL RDFHKAMESHKHYTOSETEELLSCAETDENE TLDYEEFVRRFHEPAKDIGFNVAVLLTNLSSH MPNDTRLQTFLELASSVLNYFQCHREIMG SAKRIERV YFEISESSRTQWEKPQVKESKRQFI FDVVNBGGBKERMELFVNFCEDTIERMQA QISESDLINERSAMKESEKRREFEGGRMAFF SILTVRSALFALRYNH.TLMRMLSLSKSKNOK KVKKKMTVKDMVTAFFSSYWSIFMTLHEFV ASVFROFFRIICSLLIGGSLVEGAKKIKVALHE ASVFROFFRIICSLLIGGSLVEGAKKIKVALHE HNPNAGLSDLMSNPVPMPEVQEKFQEKKAK EEKKERKEFTSEFERARGEDGEKERKAKED LTDLKELTESSDLJSDIFGLDLKREGGGYKLK EEKKERKEFTSSFERARGEDGEKERKAKED KGKQKLRQH.HTHRYGEPEVPSAFWKIIAY QQKLLNYFARNYNMRMALIVAFANNFILL FYKVSTSSVUEGKELPTRSSSEMAKVISLDSS SHRIIAVHYLEESSGYVAEPTVRIPLHITVIS FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMALALVAFANNFILL FYKVSTSSVUEGKELPTRSSSEAVKIIAY QQKLLNYFARNYVNERMELPTRIPLGITVA VARRXVNDEYGFFVGRNTSSLIAGMXAALD FSDAREKKAPKXDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT VARRXVNDEYGFFVGRNTSSLIAGMXAALD FSDAREKKAPKSDSSLOGAERTQQVAA AYGLYLRIDKGRIQCLCHSSEAGGGGERFROW BRADDRSKFVESDADELLINIPTTGIBLVC GIIMGEDDDSHFSWRLYKNIPMSFDDTER EPDQTFSLNRDLTGLELTYATKISRFSNYFHY BRADDRSKFVESDADELLINIPTTGIBLVC CNYEASANPADHRVHQVTTQTHBIS  STARL BRADDRSKFVESDADELLINIPTTGIBLVC CNYEASANPADHRVHQVTTQTHBIS HISKNFGADTTKYFYIGLRGEWFLERHEPTY CNYEASANPADHRVHQVTTQTHBIS HISKNFGADTTKYFYIGLRGEWFLERHPUT CNYEASANPADHRVHQVTQTHBIS HISKNFGADTTKYFYIGLRGEWFLARHEDYT CNYEASANPADHRVHQVTTQTHBIS HISKNFGADTTKYFYIGLRGEWFLARHEDYT TAQLRQSSLETLYVDQALFGMDTNSFVKGCL LPDQKLRLSSPVLRKAQCDQWKHSSVFSO NGELTVRAKLVPSSTRKLQEAQGEGOPKAG UPPROGRAMITYSGGCTLANDE TAGLGGGGCCAABRVQGBACSQGKLQWQKVLSSFN LWTDMTLVLH  BENKKARRICHGTFTGHTISEELQQRLDCVK QLASQPDLRDGTTYRDSBVPRSSHEDEL		İ	1				
QGECTGNQQSLAHSRLWDAVVGFLHYRAHM QMKLSQDSSQELLKELMDLGWDVMVMLLS MLEGNVVNGTGKQMVDMLVESSNYNEML KPFDPMFLKLKDLTSSDTFEKEYDPDGKGVISK RDFHKAMESHKHYTQSETEFLLSCAETIENE TLDVEBFVKRHEPAKDIGFNVAVLLTNLSHE MPNDTRLQTFLELAESVLNYFQPFLGREIMG SAKRIERTVFEISESSERTQWEKPVGENEMG GSAKRIERTVFEISESSERTQWEKPVGENEMG GSAKRIERTVFEISESSERTQWEKPVGENEMG FDVVNEGGEKEKMELFVNFCQDFTLERMG SAKRIERTVFFEISESSERTQWEKPVGENAFT GILTYRSALFALRYNLITHMRMLSLKSLKKQM KKVKKMTVKDMVTAFFSSYWSIFMTLHFV ASVSTRGFFRICSLLLGGSLVGAKKKVAELL ANMPDPTQDEVRGDGEGERKPLEAALPSED LTDLKELTEESDLLSDFGGLDLKREGGQYKLIV HNFNAGLSDLMSNPVPMPEVQEKFCGGXAK EEKEKEETKESFEKAERGGGYKEVAEL ANMPDPTQDEVRGDGEGERKPLEAALPSED LTDLKELTEESDLSDLSDFGGLDLKREGGGYKLIV HNFNAGLSDLMSNPVPMPEVQEKFCGGXAK EEKEKEERTKESFEKAERGGGYKEVAEL KOKQKLRQLHTHRYGEPEVPESAFWKKINAY QQKLLNYFARFFYMRMTLATVAFANNELL FYKVSTSSVVEGKELPTESSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVCHIVAFANNELL FYKVSTSSVVEGKELPTESSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVCHIVAFANNELL FYKVSTSSVVEGKELPTESSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVCHIVAFANNELL FYKVSTSSVVEGKELPTESSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVCHIVAFANNELL FYKVSTSSVVEGKELPTESSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVCHIVAFANNEN FCHISTORY VKRKVMDKYGEFYGRARGELLGMDKAALD FSDAREKKFKKDSSLSAVINNSDVKYQMV KLGOVFTUNSFLYLAWWINVEYQMV KLGOVFTUNSFLYLAWWINVEYQMV KLGOVFTUNSFLYLAWWINVEYQMV KLGOVFTUNSFLYLAWWINVEYQMV KLGOVFTUNSFLYLAWWINVEYQMV KLGOVFTUNSFLYLAWAL FSDARGSGGGCRCAARRYGAARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGKFRW READDRSKYFSGGGGGCCCAARRYGAARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGKFRW READDRSKYFSGGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGKFRWFW READDRSKYFSGGGGGCCCAARRYGARGSAAC GILMGGGGGCCCAARRYGARGSAAC GILMGGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGKFRWFW READDRSKYFSGGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGGCCCAARRYGARGSAAC AYGLYLRIDKGRLQCLNESSRGGGGGCCCAARRYGARGSAAC GROWN AUSTRATCHORD ATTAGATTSGARGATTARVFTLED TATAGATGGGCCCCAARRYGARGAACAC AYGLYLRIDKGRLQCLNESSRGGGGCCCCAARRY			1	1	i	1	QTGNNTTVNIIISTVDYLLRVQESISDFYWYY
QMKLSQDSSQIELLKELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVSSNNVEMIL KFPDMFLKLKDLTSSDTFKEYDPDGKGVISK RDFHKAMESHKHYTQGSTEFLLSCAETDENE TLDVEBFVKR PHEPAKDIGFNLAVLLTILSEH MPNDTRLQTFLELAESVLTNPCLGTIFEMQLAA QISESDLNERSANRESSEKERPERQGFRMAFF SILTVRSALFALRYNILTLMRMLSLSLKQM KKVKMTVXDMVTAFFSSYWSIFMTLLEFV ASVRGFFRICSLLLGGSLVEGAKKIKVAELL ANMPDPTODEVRGGDEEGEKPLEAALPSDG LTDLKERTESDLLSDIFGLDLKREGGOYKLIP HNPNAGLSDLMSNPVPMPPVQEKFOGOKA EEKEKEETKSEPEKAEGEDGKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKIIAY QQKLLNYFARNFYMMRALLFVAFAINFILL FYKVSTSSVVEGKELFTKSSERNAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSERNAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSPRINTYHNYM KLGVVFTIDNSELYLAWMT LPSCHEDDSTAMSTRANDVTGSPRINTYHOTTSPRINTYHOK BEADDESKFYKDSSLSAVTTSDRYGMW KLGVVFTIDNSELYLAWMT LPSCHEDDSTAMSTRANDVTGSPRINTYHDTSPHOLTL GIIMGEDDSIRPSEMALYKNIPQMSFDDTER EPDQTFSLNRADHVNQTCHISTS GIIMGEDDSIRPSEMALYKNIPQMSFDDTBR BEDQTFSLNRADHVNQTUTTURTS GIIMGEDDSIRPSEMALYKNIPQMSFDDTBR BEDQTFSLNRADHVNQTUTTURTS GIIMGEDDSIRPSEMALYKNIPQMSFDTBR BEDQTFSLNRADHVNQTUTTURTS GIIMGEDDSIRPSEMALYKNIPQMSFDTBR BEDQTFSLNRADHVNQTUTTURTS CHARACTAL BEGGATTAVGARACTARVFLGERGUTLAV WDFEDSTTQSFRWHPLRAKADKVEDSVPQS NGELTVAKALVLESREGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1	i	İ	i		j	SGKDVIDEQGQRNFSKAIQVAKQVFNTLTEYI
QMKLSQDSSQIELLKELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVSSNNVEMIL KFPDMFLKLKDLTSSDTFKEYDPDGKGVISK RDFHKAMESHKHYTQGSTEFLLSCAETDENE TLDVEBFVKR PHEPAKDIGFNLAVLLTILSEH MPNDTRLQTFLELAESVLTNPCLGTIFEMQLAA QISESDLNERSANRESSEKERPERQGFRMAFF SILTVRSALFALRYNILTLMRMLSLSLKQM KKVKMTVXDMVTAFFSSYWSIFMTLLEFV ASVRGFFRICSLLLGGSLVEGAKKIKVAELL ANMPDPTODEVRGGDEEGEKPLEAALPSDG LTDLKERTESDLLSDIFGLDLKREGGOYKLIP HNPNAGLSDLMSNPVPMPPVQEKFOGOKA EEKEKEETKSEPEKAEGEDGKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKIIAY QQKLLNYFARNFYMMRALLFVAFAINFILL FYKVSTSSVVEGKELFTKSSERNAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSERNAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSENAKVTSLDSS SHRIIAVHYVLESSGYMETVHSSPRINTYHNYM KLGVVFTIDNSELYLAWMT LPSCHEDDSTAMSTRANDVTGSPRINTYHOTTSPRINTYHOK BEADDESKFYKDSSLSAVTTSDRYGMW KLGVVFTIDNSELYLAWMT LPSCHEDDSTAMSTRANDVTGSPRINTYHDTSPHOLTL GIIMGEDDSIRPSEMALYKNIPQMSFDDTER EPDQTFSLNRADHVNQTCHISTS GIIMGEDDSIRPSEMALYKNIPQMSFDDTBR BEDQTFSLNRADHVNQTUTTURTS GIIMGEDDSIRPSEMALYKNIPQMSFDDTBR BEDQTFSLNRADHVNQTUTTURTS GIIMGEDDSIRPSEMALYKNIPQMSFDTBR BEDQTFSLNRADHVNQTUTTURTS GIIMGEDDSIRPSEMALYKNIPQMSFDTBR BEDQTFSLNRADHVNQTUTTURTS CHARACTAL BEGGATTAVGARACTARVFLGERGUTLAV WDFEDSTTQSFRWHPLRAKADKVEDSVPQS NGELTVAKALVLESREGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	ľ	1	1			ļ	QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM
MILEGNVVNGTIGKQMYDMLVESSNYKMEMIL KFPDMFLKLKULTSSDITFRESPIPORGVISK RDFHKAMESHKHYTQSETEFILSCAETIBNE TILDVEBEVKRPHEPAKDIGFONAVILTIN.SEH MPNDTRLQTIFLELAESVLNYFQPPLGRIEMG SAKRIERVYFEISESSRTQWERFQVKESKQFI FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA QISESDLNERSANKESEKERPEQGFMAFF SILTVRSALFALRYNHLTIMRMLSLKSLKKQM KKVKKMTVKDMYTAFFSSYWSIFMTLLHFV ASVERGFRICSLLGGSLVGAKKKVAELL ANMPPPTQDEVRGDGEEGRKPLEAALPSED LTDLKELTESDLLSDIFGLDLKRRGGGYKLIP HNFNAGLSDLMSNPVRMFEVQEKFQEVAK EEKEKEETKSEPEKAEGBDGKEKERAKED KGKQKLRQLHHRYGGEPYPESAFWKKILAY QQKLLNYFARNFYMRMLALFVAFANFILL FYKVSTSVVEGKELPTISSESNAKVISLDSS SHRIIAVHVVLEESSGYMEPTVRLIPHTVISF FCIIGYYCLKVPLVFKREKEVARKLEFDLTVISF FCIIGYYCLKVPLVFKREKEVARKLEFDLTVISF FCIIGYYCLKVPLVFKREKEVARKLEFDLTVISF FCIIGYYCLKVPLVFKREKEVARKLEFDLTVISF FCIIGYYCLKVPLVFKREKEGGAERTYQVAA LJESHGISHGGGGCRCAERVGAARGSAC AYGLYLRIDKGRIQCUNESREGGGGEFTQQVAA LJESHGISHGGGGCRCAERVGAARGSAC AYGLYLRIDKGRIQCUNESREGGGREFTQVAK GIIMGEDDDSIFPSEMILSTUPMSHFDITGIFUKLK GIIMGEDDDSIFPSEMILSTUPMSHFDTIGHVILK GIIMGEDDDSIFPSEMILSTUPMSHFDTIGHVILK GIIMGEDDSIFPSEMILSTUPMSHFDTIGHVILK GIIMGEDDSIFPSEMILSTUPMSHFDTIFTGIFTGIFTGIFTGIFTGIFTGIFTGIFTGIFTGIFT	1		1	1		1	OMKLSQDSSQIELLKELMDLQKDMVVMLLS
RFPDMFLKLKDLTSSDTFKEYDPDGKGVISK   RDFHKAMESHKHTYGSETELLSCAETIDENE   TLDVEBFVKRHEPAKDIGFNVAVLINLSEH   MPNDTRLOTTELLASSVLVHYGOPFLGRIEMG   SAKRIERVYFEISESSRTÖWEKPQVKESKQOFI   PDVVNEGGEKEKMELFVNPCEDTIFEMQLAA   QISESDLNERSANKESESEKERPEEQGFRMAFF   SILTVRSALFALRYNILTIMRMLSLSJKKOM   KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV   ASVFRGFFRICSLLLGGSLVEGAKKIKVAELL   ANMPDPTQDEVRGGDEGEKPLEADLPSED   LTDLKELTESSDLLSDFGLDLKREGGOYKLIF   HNPNAGLSJUMSNPVPMPEVQEKFOGKAK   EEREKEFKEFTKSEPPKAEGBOBKEKEKAED   KOKQKLRQLHTHRYGEPPVESAFWKKILAY   QQKLLNYPARNFYNMRMLALFVAFANFILL   FYKVSTSSVVEGKELPTRSSSENAKVISLOSS   SHRILAVHVVLEESSGYMEVTVALIPHHTVISF   FCIIGYYCLKVPLVIFKREKEVARLEFDGLYT   TEQPSEDDIKGQWDRLVINTGSPPNYWDKF   VKRVMDKYGEFYGBRISELLGMDKAALD   FSDAREKKKPKDSSLSAVLNSIDVKYQMV   KLGVVFTDNSFLYLAWMT   LPELNGRGAGLRAFESEEGGGAERTQOVAA   AVGLVLRIDKGRUQCLNESSRGGGRGVFFW   ERADDRSKYFVSDADDELLFINTFTGHVELK   GIIMGEDDSHPSEMRLYKNIFQMSPDDTR   EPDQFFSLARGARGACAC   AVGLVLRIDKGRLQCLNESSRGGGAERTOVFFW   ERADDRSKYFVSDADDELLFINTFTGHVELK   GIIMGEDDSHPSEMRLYKNIFQMSPDDTR   EPDQFFSLARGAPHYVETGLRGEWTELRREPYT   CYNFASANPADHRYNGYCTHIS   HISKNFGADTTKYFYIGLRGEWTELRREPYT   CYNFASANPADHRYNGYTCHIS   HISKNFGADTTKYFYIGLRGEWTELRREPYT   CYNFASANPADHRYNGYTCHIS   EPDQFFSLARGARGFYVKTVILLD   RSSQKRRTTGVGRNTVPFTFGETLKYQVAPP   QLVTRQLQVSVWHLGTLARRVFLGEVILPA   WDPEDSTTQSFR WHEIRAKADKYEDSVPQS   NGELTVRAKLVLPSRTRKLQBAQBGTNOPSL   HGQLCLVVLGAKNPVKTVLGEVILPA   WDPEDSTTQSFR WHEIRAKADKYEDSVPQS   NGELTVRAKLVLPSRTRKLQBAQBGTNOPSL   HGQLCLVLGAKNPVKTVLGEVILPA   WDPEDSTTQSFR WHEIRAKADKYEDSVPQS   NGELTVRAKLVLPSRTRKLQBAQBGTNOPSL   HGQLCLVLGAKNPVKTVLGEVILPA   WDPEDSTTQSFR WHEIRAKADKYEDSVPQS   NGELTVRAKLVLPSRTRKLQBAQBGTNOPSL   HGQLCLVLGAKNPVKTVLGEVILPA   WDPEDSTTQSFR WHEIRAKADKYEDSVPQS   TAQLRQSKLLTVWDQALJFGMMDRLLGGT   LPDQKLRK KSPVLRKQACPQWKNSFNS   LPDQKLRK KSPVLRKQACPQWKNSFNS   LPDQKLRK KSPVLRKQACPQWKNSFNS   LPDQKLRK KSPVLRKQACPQWKNSFNS   LPDQKLRK KSPVLRKGACPQWKNSFNS   LPDQKLRK KSPVLRKGACPQWKNSFNS   LPDQKLRK KSPVLRKGACPQWKNSFNS   LPDQKLRK KSPVLRKGACPQWKNSFNS   LPDQKLRKGACPQWKNSFNS   LPDQKLRKGACPQWKNSFNS   LPDQK	1	1					MLEGNVVNGTIGKOMVDMLVESSNNVEMIL
ROFHIKAMESHKHYTQSETEFLLSCAETDONE TLDYSEPVKRHPPAKDIGNNAVLITNLSEH MPNDTRLQTFLELAESVLNYTQPFLGRIEMG SAKRIERVYFEISESSRTQWEKPQVKESKRQFI FDVVNEGGEKEKMELPVNFCEDTIFEMQLAA QISESDLNERSANKEESEKERPEGGFMAFF SILTYRSALFALTYNILTIMEMLSLKSLKKQM KEVKKMITVKDMYTAFFSSYWSIFMTLLHFVY ASVFRGIFFRICSLLLGSSLVEGAKKIKVAELL ANMPDFTQDEVRGDGEEGERKPLBAALPSED LTDLKELTEESDLLSDFELKREGGYKLIF HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK EEEKEKEETKSEPKR-AGGDGEKEKAKELD KOKQKLRQHTHRYGEPEVPESAFWKNIAY QOKLLNYFARNFYNMRMLALFYAFANFILL FYKVSTSSVYGKELPTRSSSENAKVTSLDSS SIRRIAVHYVLEESSGYMEPTVRILPLHTVISF PCIGTYYCLK VPLVIFKREKEVARKLEFDGLYI TEOPSEDDIKOQWDRLVNTOSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKEPKKOSSLSAVLNSIDVKYQMV KLGVVFTDNSFLYLAWYMI LPENNGGRGGERCAERSEGGGGERTQQVAA LPLSIGHSHGGGGCCAAERVGAARGSAAC AYGLYLRDKGRJCCHXNPQOMSPDDTER EPDQTFSLNRDLTGELEYATKISRESSNYHLS GIIMGEBODSHESENKENSEGSGGRGYERVW ERADDRSKFVESDADEELLFNIPFTGHVKLK GIIMGEDDSDHESENKENKENGPOMSPDTER EPDQTFSLNRDLTGELEYATKISRESSNYHLSI HISKNFGADTTKVFYIGLRGEWTERRHEVTI S54  1904 A 4200 1 961 GIPCTEMGRIFDNANVTGELEFAHTCFKTHSL UNYEASANAPAHEKVHQVTQFTHFIS CINKACKNLAYGEEKKKCNPYVKTYLLPD RSSQGKKTGVORNTVDPTFGETILKYQVAP QLVTRQLOVSVHLGTARRVFJGEVIPLA' WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAGGTOQPSL HGQLCLVVLGAKNLPVPRDGTLNSEVKGCLI LPDQKLRLKSFVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVVQLAGKMDRLLGGT RLGSKGDTAVGGDACSQSKLQWGKVLSSFN LWTDMTLVLH KENKKARNLRNNQSRSRSDGGSETLPODH NHHENERWQQERLHREEAYYQFIPELINGE DYRLMRDSNLLGTTGGITSSELQQRLDVKE QLASQPLLRGGTNYRDSEVPRESSHEDSLLE WINTFRITGGATATSGQNGNOTWRAVSSTNI NNGEFRFSLEHIVNIEENRGFEHIEDDYTDILL DSINBHTARNQGRSTSPVARRTRSGTSVNFI NNGEFRFSLEHIVNIEENRGFEHIEDDYTDILL DSINBHTARNQGRSTSPVARRTRSGTSVNFI	]	1	1	1	1	1	
TILDYEEFVKRHIPEADIGHNAVLITILISEH MPMOTIKLOTTELAESUNYTOPPIGRIEIMG SAKRIERVYTEISESSRTQWEKPQVKESKROPI FDVVNEGGEKEMELFVNFCEDTTEMQLAA QISESDLNERSKANKEESEKEPEEQGRMAFF SILTYRSALFALRYNILTLMRMLSLSSLKQM KKYKKMYKOMYTAFSWSIFMTLLHFV ASVFRGFFRIICSLLGGSLVEGAKKIVAELL ANNEDPTODEVRGOGEGERKPLEAALPSED LTDLKELTESDLLSDIFGLDLKREGGQYKLIP HHPNAGLSDLMSNPVPMEVQEKPÖRÇAK EEEKEEKEEKEEFEKERFEKAEGGGEKPLEAALPSED LTDLKELTESDLLSDIFGLDLKREGGGYKLIP HHPNAGLSDLMSNPVPMEVQEKPÖRÇAK EEEKEEKEEKESPEKAEGGGGKEEKARED KOKQKLRQLHTHRYGEPEVPESAFWKHIAY QOKLLNYFARNFYNMRMIALFVAFANFILL FYKVSTSSVVEGKELPTRSSSENAKVISLOSS SIRHIAAVHYLEESSGYMEPTVRILPILHTVISF PCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEOPSEDDIKGOWDRLVINTQSFPNYMDKAF VKRKVMKYGEFYGERNSELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW LIGVVFITDNSFLYLAWREN LIGVVFITDNSFLYLAWREN VKRKVMKYGEFYGERNSELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW LIGVVFITDNSFLYLAWREN LIGVVFITDNSFLYLAWREN ERADDRSKFVESDADEELLFNIFTGHVKLK GIIIMGEDDSHPSISMLRANPOMSPDDTER ERDOTFSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWETLERHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTFSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWETLERHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTFSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWETLERHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTFSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWETLERHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTTSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWTELRRHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTTSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWTELRRHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTTSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWTELRRHEVTI CNYEASANPADHRVHQVTPOTHFIS EDOTTSLNRDLTGELEYATKISRFSNYYHLS HISKNFGADTTKVFYLGEWTELRRHEVTI CNYEASANPADHRVHQVTPOTHFIS  554  1904  A 4200  1 961  GIPCTEMGNFDNANVTGELEFATHYCKTCHSL GIPCTEMGNFDNANVTGELEFATHYCKTCHSUBL WDFFTALRRUCHLENGEVTELIBLA WDFEDSTTOSFRWHPLRAKADKYFLGEVIPLA WDFEDSTTOSFRWHPLRAKADKYFLGEVIPLA WDFEDSTTOSFRWHPLRAKADKYFLGENDE DYRLMRODNLLGTTGGETSELQORLDUKE QLASQPILRTGTNYRDSEVPRESSHEDSLLE WLNTFRITGGNATARSQORNOTWAVSSTNI NNGEFRESLEHIVNHENRGFEHIEGDYTDIPL DSNEMHTANRQQRSTSPVARTRSQTSVNF	l	1		1	Į.		1
MPNDTRLQTFLELAESVLNYFQPFLGRIEMG SAKRIERVYFISIESSRYDKPQVESSRQFI PDVVNEGGEKEKMELFVNFCEDTIFEMQLAA QISESDLNERSANKEESEKERPEGGFRMAFF SILTVRSALFALRYNILTIMEMLSLKSLKKQM KEVKKMITVKDMYTAFFSSYWSIFMTLIFIY ASVFRGIFFRICSLLLGSSLVEGAKKIKVAELL ANMPDPTQDEVRGDGEGERFPLEAALPSED LTDLKELTEESDLLSDIFGLDLKREGGQYKLIF HNPNAGLSDLMSNPYPMEVQEKPGEÇGKAK EEEKEKEETKSEPEKAGEDGEKEKAKELD KOKQKLKQLHTHRYGPEPVPSSAFWKKILAY QOKLLNYFARNFYNMRMLALFYGANFILL FYKVSTSSVVEGKELFIPGLYIT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW KLGVVFIDNSFLYLAWYMT TEQPSEDDIKGQWDRLVINSDVKYQMW ERADDRSKYVESDADEBLLFNIPFTGHIVKLK GIIMGEDDDSHPSENRLVKNIPQNSFDTER EPPQTFSLNRDLTGELFYAFTGSSFORYFEYM ERADDRSKYVESDADEBLLFNIPFTGHIVKLK GIIMGEDDDSHPSENRLVKNIPQNSFDDTER EPPQTFSLNRDLTGELFYAFTGSSFORYFEYMT ERADDRSKYVESDADHVHQVTTQTHIFIS EICKACKNLAYGEKKKKCNPYKTYLTLLPD RSSQGKKKTOVQRNTYGEBFAHYCFCKTHSL EICKACKNLAYGEKKKKCNPYKTYLTLLPD RSSQGKKTOVQRNTYGEBFAHYCFORYPG NGELTVRAKLVLPSSTRIKLJGBAGEGTDQPSL HGQLCLVVLGAKNLPVRPGTILNSFVKGCL LPPQOKLRKSPVLRKAQAPQWKHSFVSG TPAQLRQSSLELTVWQALFGAMDRLLGGT RLGSKGDTAVGGDACSQSKLQWKVLSSPN NGELTVRAKLVLPSSTRIKLJGBAGEGTLPQDH KENKKARNILRNNQSRSRSDGGSETLPQDH NHENERERWQQELHKREAYYQFINELNDE DYRLMRDINLLGTPGEITSELQQRLDGVKE QLASQPDLRGGNTYRDSEVPRSSHEDSLLE WLNTFRTKGATARSGGRONGTWAVSRTNI NNGEFRSSLEHTVNHENRGFEHIEGDYTDIFL DSNEMHTANRQQRSTLFRANGSTNINN				1		1	
SAKRIER VYEISESSKTQWEKPQVESKROPT POVVNEGGEKEMELL-PNFCEDTTEMQLAA QISESDL NERSANKESSEKERPEGOPRMAFF SILTVRSALFALRYNILTILMRMLSLKSLKKOM KKYKKNIYCKDWYTAFFSYWSIFMILLHPV ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL ANNPDPTODEVRGOGEGERKPLEAALPSED LTDLKELTESDLLSDIFGLDLKREGGGYKLIP HNENAGLSDLMSNPVPMEVQEKPGEGVAK EEEKEEKEETKSEPEKAEGDGEKEK AKED KOKQKLRQHTIRRYGEPEVPESAFWKKILAY QOKLLNYFARNFYNMENMALLFVAFANFILL FYKVSTSSVVBGKELPTRSSSENAKVSLIDLS SIRIIAAVHYULESSGYMEPTVRILPILITVISF PCIIGYYCLKVPLVIFKREKEVARKLEFDOLTY TEOPSEDDIKOQWDRLVTMDSFPNYWDKF VORKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRPKKDSSLSAVLNSIDVKYQMV VORKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRPKKDSSLSAVLNSIDVKYQMV VORKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRPKKDSSLSAVLNSIDVKYQMV VORKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRPKKDSSLSAVLNSIDVKYQMV VORKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRPKCDSSLSAVLNSIDVKYQMV LLPLSHGIGSGGCGCAABRVQGARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKFW ERADDRSKFVESDADEELLFNIPFTGHVKLK GIIMGEDDDSHPISEMLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRSNYYHLS HISKNFGADTTKVFVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS EPDQTFSLNRDLTGELEYATKISRSNYYHLS HISKNFGADTTKVFVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS EPDQTFSLNRDLTGELEYATKISRSNYYHLS HISKNFGADTTKVFVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS EPDQTFSLNRDLTGELEYATKISRSNYYHLS GPPCTEMGFFDBNAVTGEBFFAHYVFCKTHSL EICIKACKNLAYGEEKKKCCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFGEITLKYQVAP QLATRGLQVSVWHLGTLARRVFLGEVIPLA* WDFFDSTTGSFFSWHPLRAKADKYEDSVPQS NGELTYRAKLVLPFSRTU,GAQGFOTDQPSL HGQLCLVVLGAKKLPRPDGTLNSFVKGCT LPPQQKLRLKSPVLRKQACPQWKHSSYPSG TPAQLRQSSLELTVVDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWKVLSSFN TPAQLRQSSLELTVVDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWKVLSSFN TPAQLRQSSLELTVVDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWKVLSSFN TPAQLRQSSLELTVVDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWKVLSSFN TPAQLRQSSLELTVVDQALFGMNDRLLGGT RLGSKGDTAVGGDACSGSKLQWKVLSSFN TPAQLRQSSCGRRTGVTRAVSSTNEN NNGEFRFSLEHIVVNIENRGFEHIGEDYTDJEL DYRLMRDHNLLGTPGEITSELQORLDGVKN NNGEFRSSLEHIVVNIENRGFEHIGEDYTDJEL DSNRDHTARRQQFTATSPVARTRSGTSVFF	1.				!	1	MONDALD CALLED VICTOR AND THAT AND THE COURT
FDVVNEGGEKEMELFVNFCEDTITEMQLAA   QISSEDLERSANKESEKERPEQOPRMAFF   SILTVRSALFALRYNILTIMRMLSLXSLKKQM   KKVKKMTVKDMYTAFFSSYWSIFWITLLHFV   ASVFROFFRIICSLLGGSLVEGAKKKVAELL   ANMPDPTQDEVRGDGEGGEKKLVAELL   ANMPDPTQDEVRGDGEGGEKFKLPAALPSDE   LTDLKELTESDLJSDIFGLDLKREGGQYKLIF   HNPNAGLSDLMSNPVPMPEVQEKFQECKAK   EEEKEEKEEKETKSEPKAGDEKEEKKAKED   KGKQKLRQLHTHRYGEPEVPESAFWKRIIAY   QQKLLNYFARNFYYMMRMLALTVAFAINFILL   FYXVSTSSVVEGKELFTRSSENAKVISLDSS   SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF   PCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI   TEQPSEDDIKQQWDRLVINTQSFPNNYWDKF   VKRKVMDKYGEFYGRDRISELIGMDKAALD   FSDAREKKPKKDSSLSAVLNSIDWKYQMW   KLGVVFTDNSFLYLAWYMT   VKRKVMDKYGEFYGRDRISELIGMDKAALD   FSDAREKKPKKDSSLSAVLNSIDWKYQMW   KLGVVFTDNSFLYLAWYMT   LPELNGRGAGLRRAEPSERGGGAERTYQVAA   AVGLYLRIDKGRLQCLNESSEGSGROVFKFW   ERADDDSKFYVESADAEELLFNIPFFGIVKLK   GIIMGEDDDSHPSEMRLYKNIPQMSPDDTER   EPDOTTSLNRDLTGGELYATKISRSFNVYHLSI   HISKNFGADTTKVFYJGLRGEWTELRRHEVTI   CNYEASANPADHRVHQVTPQTHTHS   EPDOTTSLNRDLTGGEFAIHYCKTHSL   EICIKACKNLAYGEEKKKCNPYVKTYLLPD   RSSGKRTGVQFNDFTFGETLKYQVAP   QLVTRQLQVSWHLGTLARRVPLGEVIBLA*   WDFEDSTTQSFRWHPLRAKADKYEDSVPQS   NGELTVRAKLVLPSTPGTLNSTVRGCT   LPDQQKLRLSSPVLRKQACPQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACCQWKHSFVPSG   TPAQLRQSSLETVWDQALFGMNDRLLGGT   RLGSKGDTAVGGACSQSCELLPQDP   HGQLCLVVLGANNLFRREGSHEDKLE   WINTFRRTGNATRSGQNGRQTWRAVSRTM   NHEBERRVQCEHKBEAYYQFINEINDE   DYRLMRDHNLLGTPGGITSELLQSLLDDFNDDTTANRCORSTSSYVARRTSGONNOTWAVSRTM   NHEBERRVQCEHGEDCYTDIFL   DSNRDHTANRCORSTSSYVARRTSGONNOTWAVSRTM   NNGEFRISLEHVNHENGGEHGEDYTDIFL   DSNRDHTANRCORSTSSYVARRTSGONNOTWAVSRTM   NNGEFRISLEHVNHENGGEHGEDYTDIFL   DSNRDHTANRCORSTSSYVARRTSGONNOTWAVSRTM   NNGEFRISLEHVNHENGGEHGED	İ		1	1	1	1	O TENED MESICE COLLONIES A PHILAL LOCKERNO
OISESDLNERSANKEESEKERPEEQOPRMAFF SILTVRSALFALRYNILTLMRMLSLXSLKKOM KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV ASVFROFFRIICSLLIGGSLVEGAKKIKVAELL ANMPDPTODEVRGOEGEGERFLEAALPSED LTDLKELTEESDLISDIFGLDLKREGGQYKLIP HNPNAGLSDLMSNPVPMPEVQEKFQEQYKLIP HNPNAGLSDLMSNPVPMPEVQEKFQEQYKLIP HNPNAGLSDLMSNPVPMPEVQEKFQEQYKLIP GOKLINTFARNFYMMRMLALFVAFANIFILL FYXVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIAVHYVLEESSGYMMRMALFVAFANIFILL FYXVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLIVFREKEVARKLEFDGLYI TEQPSEDDIKGQWDRILVINTOSFPNNYWMC VGREYMDKYGGFFGRRISELLGMDKAALD FSDAREKKFKKDSSLSAVLNSDVKYQMW KLGVVFTDNSFLYLAWYMT  1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAERYGARAGSAAC AYGLYLRDKGRILQCLNESSEGSGRGVFRW ERADIDRSKYVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPPQTTSLNRDLTGGLEVATKISRSNVYHLSI HISKNFGADTTKVFYYIGLRGEWTELRRIEVTI CNYEASANPADHRVQVTPQTHFIST  554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFALHYCFKTHSL EICIKACKNLAYGEKKKCNPYVKTYLLPD RSSQKKRTGVQRNTVDFTTGETLKYQVAPA QLVTRQLQVSVWHLGTLARAVFLGEVIPLAT WDFEDSTTQSFRWHPLRAKADKYEGSVPGS UPGCKLRLSSPYLKQACCPQWKHSFVFSG TPAQLRQSSLEITVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACQSKLQWGKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNIRMNQSRSSDGGSEETLPQDH NHEHERRWQGERLHREEAYYQFINEINDE DYRLMRDHNLLGTPGITISBSEUGLIGE UNTTRRTGNATRSGQNONQTWRAVSRTNI NNGEFRISLEIHVNHENGIFEHGEDYTDIFL DSNRDHNLLGTPGBITSBSEVPRESSHEBSILE WLNTFRRTGNATRSGQNONQTWRAVSRTNI NNGEFRISLEIHVNHENGIFEHGEDYTDIFL DSNRDHTANROGRSTSSPVARRTRSGTSVHFI NNGEFRISLEIHVNHENGIFEHGEDYTDIFL DSNRDHTANROGRSTSSPVARRTRSGTSVHFI NNGEFRISLEIHVNHENGIFEHGEDYTDIFL DSNRDHTANROGRSTSSPVARRTRSGTSVHFI			}	}	1	1	
SILTVRSALFALRYNILTIEMRISLKSLKKQM KKWKKMTVKDMYJAFRSYWSIFNTLHFV ASVFRGFFRICSLLLGGSLVEGAKKIKVAELL ANMPDPTQDEVRGDGEEGBERPLEAALPSED LTDLKELTESSDLLSDIFGLDLKREGGQYKLLF HNPNAGLSDLMSNPVPMBEVDGEKFQEGKAK EEKEKEKESTSEPBAGUQEKEFQEGKAK EEKEKEKEKETKSEPBAGUGEKERGEKAKED KGKQKLRQLHTHRYGEPPVPESAFWKKILAY QQKLLNYFARNFYNMRMLALFVAFANFILL FYXVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRLPILHTVISF PGIGYYCLKVPLVFREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKFKKDSSLSAVLNSIDVKYQMV KLGYVFTDNSFLYLAWYMT LPELNGRAGGLRRAEPSERGGGAERTQQVAA LPI.SHGHISHGGGGCGCAAERYGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKFVESDADEELLPNIPFTGHVKLK GIIMGEDDDSHPSEMRLYKNIPQMSPDDTER EPDOTTSLNRDLIGGELEYATKISRFSNYHLSI HISKNGADTTKVFYIGLRGEWTELRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGELEFAHYCKTHSL EICIKACKNLAYGEEKKKCNPYVKTYLLPD RSSGKRKTGVGNYDDTFQCFLKYQVAP QLVTRQLQVSWHLGTLARRVFLGEVIPLA* WDFEDSTTQSFRWHPLRAKADKYFDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTINSFVKGCT LPDQKIRLKSPYLRQACPQWKHSFVPSG TPAQLRQSSLELTVWDQALFGMNDRLLGTT RLGSKGDTAVGGDACSQSKLQWGKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSSRSDGSEETLPQDH NHEHERRWQCERLHREEAYYQTRIENDE DYRLMEDHNLLGTPGGITSECLQGRLDGVKE QLASQPDLRDGTNYRDSSVPRESSHEDSLE WLNTTRRTGNATTSGQNGNQTWRAVSRTNE NNGEFRISLEHVNHENGFEHIGEDYTDLD DYRLMEDHNLLGTPGGTSSEHEDSLLE WLNTTRRTGNATTSGQNGNQTWRAVSRTNE NNGEFRISLEHVNHENGFEHIGEDYTDLD DSNRDHTANRQORSTSPARRTRSGTSVNFI	1	1		1			
KKVKMITVKDMYTAFFSSYWSIFMTLLHFV ASVPRGFRITICISICLIGGSLVEGAKKIVABLL ANMPDPTQDEVRGDGEGERKPLEAALPSED LTDLKELTEESDLLSDIFGDLKREGGVYKLIF HNPNAGLSDLMSNPVPMPEVQEKFQEÇKAK EEEKEKEETKSEPEKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGPEPVPESAFWKKIJAY QQKLLNYFARNFYNMRMLALFVAFAINFILL FYKVSTSSVVBGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESGYMEPTVREIDLHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKYMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRPKKDSLSAAVINSIDVKYQMW KLGVVFTDNSFLYLAWYMT  553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQVAAA LPLSHGHSHGGGGCCAARVVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKFVESDADELLFINTPTIGHVKLK GIIIMGEDDDSHPSEMMELYNNPGMSFDOTER EPDQTFSLNRDLTGELEVATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYGASANPABHRVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGBIEPAHHYCFKTHSL EICIKACKNLAYBERKKKKCNPYKTYLLPO RSSQKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIBLA* WDFEDSTTOSFREKKKKCNPYKTYLLPO RSSQKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIBLA* WDFEDSTTOSFREKKKKCNPYKYTYLLPO RSSGGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIBLA* WDFEDSTTOSFREKKKKCNPYKTYLLPO RSSGGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIBLA* WDFEDSTTOSFREWKLRKACNPYKTYLLFO RLGGGGTAVGASSLELTVWDQALFGMNDRLLGGT RLGSGGDTAVGGSSLELTVWDQALFGMNDRLLGGT RLGSGGDTAVGGSSLELTVWDQALFGMNDRLLGGT RLGSGGDTAVGGSSLELTVWDQALFGMNDRLLGGT RLGSGGTAVGQCRLHREEAYYQFINELNDE S55 1905 A 4211 331 2419 KENKKARNLRNNQSRSESGGSEETLPQDH NHHENBERWQCGCRLHREEAYYQFINELNDE QLASQPDLRGGGTPYRDSSHPDSLLS WLNTFRTGNATRSQQNONQTWRAVSTNI NNGEFRFSLEHIVNHEINRGFEIHGEDYTDIPL DSNRDHTANRQGRSTSSPHARRTRSGTSVNYNT NNGEFRFSLEHIVNHEINRGFEIHGEDYTDIPL DSNRDHTANRQGRSTSSPHARRTRSGTSVNYNT	1	1	Į.		i	1	QISESDLNERSANKEESEKERPEEQGPRMAFF
ASVPRGFFRICSLLLGGSLVEGAKKIKVAELL ANMPDPTQDEVRODGEEGRKRICAALPSED LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP HNPMAGLSDLMSNPPVPMPEVQEKFQEQKAK EEEKEKEETKSEPEKAEGEDGEKEEKAKEND KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYMMRMLALFVAFAINFILL FYKVSTSSVVEGKELPTRSSSEMAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRLIPLHTVISF PCIIGYYCLKVPLKFREKEVARKLEFDGLYJ TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRYMDKYGEFYGRDRISELLGMDKAALD FSDAREKKEPKKEKKEAKLEFDGLYJ TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKYMDKYGEFYGRDRISELLGMDKAALD FSDAREKKEPKKESLSAVINSIDDVKYQMW KLGVVFTDNSFLYLAWYMT  LPELNGGGAGRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAABEN/QAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADJDRSKFVESDADEELLFNIPFGGHVELK GIIIMGEDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554  1904 A 4200 1 961 GIPCTEMGNFDNANVTGELEFAHYCFKTHSL EICKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQQKRKTGVQRNTVDPTQETILKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLAK. WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTILKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLAK. UPDQKLRLKSPVLRXQACPQWKHSFYSGG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSSGDTAVGGDACSQSKLQWGKVLSSPN LWTDMTLVH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHINLLGTPGEITSELQQRLDGVKE QLASGQDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNI NNGEFRFSLEHIVNIEENREFEHGEDTUPILL DSNRDHTANRQGRSTSPVARRTRSGTSNYNT NNGEFRFSLEHIVNIEENREFEHGEDTUPILL DDSNRDHTANRQGRSTSPVARRTRSGTSNYNT	1 .	1	1	1 .			SILTVRSALFALRYNILTLMRMLSLKSLKKQM
ANMPDPTQDEVRGOGEGERKPLEAALPSED LTDLKELTEESDLISDIFGLDLKREGGYYKLIW HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK EEEKEKEETKSEPEKAGEGOTKLIW OKOKLRQLHTHENFGEBUGEKEEKAKED KGKOKLRQLHTHENFGEBUGEKEEKAKED KGKOKLRQLHTHENFGEBUGEKEEKAKED KGKOKLRQLHTHENFGEBUGEKEEKAKED KGKOKLRQLHTHENFGEBUGEKEEKAKED KGKOKLRQLHTHENFGEBUGEKEEKAKED KGKOKLRQLHTHENFAGEBUGEKEEKAKED VIRVARANFVNIMAMALFVAFANFILL FYKVSTSSVVEGKELPTRSSSENAKVISIDLYTISDIF FCIIGYYCLKVPLVIFKEEKEVARKLEPDGLYI TEQPSEDDIKGQMELVINTQSPFINYPDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFIDNSFLYLAWYMI  LPELNGRGAGLRRAEPSEGGAERTQQVAA LPLSHGSHGHGGGCCCAABRVQAARGSAAC AYGLYLRIDKGRLQCLNESREGGGROFKFW ERADDRSKFVESDADEELLFINFFTIGHVKLK GIIIMGEDDDSHFSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEVATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRHEVTI CNYEASANPADHEVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMONFDNANVTGEBERAHFYCFKTHSL EICKACKNLAYGEKKKKCNPVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETILKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLAY WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGKNLYPRDGTLNSFYKGCLJ LPDQKKLRKSPVLRKQACPQWKHSFVPSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSGGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVH  555 1905 A 4211 331 2419 KENKKARNLRMQSRSSSDGSEETLPQDH NHHENERRWQQCRLHREEAYYQFINELNDE QLASQPOLRDGTNYRDSEVPRESSHDDSLLE WLNTFRRTGNATRSQQNGNQTWRAVSRTNI NNGEFRFSLEHIVNIENRGFEHGEDTUDIFL DSNRDHTANROQRSTSSPVARRTRSQTSNVMF NNGEFRFSLEHIVNIENRGFEHGEDTUDIFL DSNRDHTANROQRSTSSPVARRTRSGTSNVMF		l	1	<b>1</b>			KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
ANMPDPTQDEVRGOGEGERKPLEAALPSED LTDLKELTEESDLLSDIFGLDLKREGGYYKLIW HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK EEEKEKEETKSEPEKAGEDGEKEEKAKED KGKQKLRQLHTHEKFGEBUGEKEEKAKED KGKQKLRQLHTHEKFGEBUGEKEEKAKED KGKQKLRQLHTHEKFGEBUGEKEEKAKED KGKQKLRQLHTHEKFGEBUGEKEEKAKED KGKQKLRQLHTHEKFGEBUGEKEEKAKED KGKQKLRQLHTHEKFGEBUGEKEEKAKED KGKQKLRQLHTHEKFTGEBUFYTELPFLHTVISF FCIIGYYCLKVFLVIFKEEKEVARKLEFDGLYI TEQPSEDDIKGQMELVINTQSFPINYPDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFIDNSFLYLAWYMT  LPELNGRGAGLRRAEPSEGGAERTQQVAA LPLSHGSHGHGGGCCCAABRVQAARGSAAC AYGLYLRIDKGRLQCLNESREGGGROFKFW ERADDRSKFVESDADEELLFINFFTIGHVKLK GIIIMGEDDDSHFSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKVFYIGLRGEWTELRHEVTI CNYEASANPADHEVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMONFDNANVTGEBERAHFVTY CNYEASANPADHEVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMONFDNANVTGEBERAHFVTY CNYEASANPADHEVHQVTPQTHFIS  555 1905 A 4211 331 2419 KENKKRNPLVRPDGTLNSFYKGCL, LPDQQKLRLKSPVLRQACPQWKHSFVPSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVH  KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQGRLHREBAYYQFINELMDE DYRLMRDHINLLGTPGEITSELQQRLDGVKE QLASQPDLRDGTNYRDSSVPRESSHDDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNI NNGEFRFSLEHIVNIEENRGFEHGEDTUDIFL DSNRDHTANRQORSTSPVARRTRSGTSVNMF	1	1	1	1	l	1	ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL
LTDLKELTESSDLLSDIFGLDLKREGGYKLIP   HNPNAGLSDLMSNPVPMPEVQEKFQEDKAK   EEEKEKEETESTEKAEGEDGEKEKAKED   KGKQKLRQLHTHRYQEPEVPESAFWKKIIAY   QQKLNYFARNFYMMRILALFVAFANFILL   FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS   SHRIIAVHYVLEESSGYMEPTVRILPLHTVISF   FCIIGYYCLKVPLVFKREKEVARRLEFDGLYI   TEQPSEDDIKGQWDRLVINTQSPPNNYWDKF   VKRKVMDKYGEFYGRDRISELLGMDKAALD   FSDAREKKRKPKDSSLSAVINSIDVKYQMW   KLGVVFTDNSFLYLAWYMT   LPELNGRAGGLRAEPSERGGAERTQQVAA   LPLSHGHSHGGGGCRCAAERVGAARGSAAC   AYGLYLRIDKGRLQCLDESREGSGRGVFKPW   ERADDRSKFYESDADEELLFNIPFTIGHVKLK   GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER   EPDQTTSLNRDLTGELEYATKISRFSNVYHLSI   HISKNFGADTTKVPYIGLRGEWTELRRIEVTI   CNYEASANPADHRVHQVTPQTHFIS   STATEM			i	ł	1		ANMPDPTQDEVRGDGEEGERKPLEAALPSED
HNPNAGLSDLMSNPVPMEVQEKFQEQKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYMMRIALFVAFANFIIL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVFKREKEVARKLEPDGLYT TEQPSEDDIKGQWDRLVNTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKRFKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT  553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCCAAERVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKFVESDADEELIFNIPFTGHVKLK GIILMGEDDSHPSEMFLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRIEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGELEFAHLYCFKTHSL EICIKACKNLAYGEKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQEILXQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIPLAX WDFEDSTTQSFRWHEJRAKADKYEDSVPQS NGELTVRAKLVPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGA TPAQLROSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH KENKKARNICRIMQSRSRSDGGSEETLFQDH NHHENERR WQQERLHREAYYQFINELNDE DYRLMRDHNLLGTTPGEITSELQQRLDGVKE QLASQPDLRAGGTNYRDSEVPRESSHEDSLLE WLNTFRTGNATRSGQNGNQTWRAVSRTNI NNGEFRISLEIHVHLENGFEITGEBGLLGEVTDIPLA DSNRDHTANRQQSSTSSVARRTRSGTSSNFT		İ		}	j	1	LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP
EEEKEEKETKSEPBKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGEPBVPESAFWKKIIAY QQKLLNYFARNFYMMRMLALFVAFAINFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSPPNNYMDKF VKRKVMDKYGEFYGRDISELLGMDKAALD FSDAREKKRPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT  553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSIGGGCRCAAERIVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADIDRSKFVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGGLEFATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHPIS  554 1904 A 4200 1 961 GIPCTEMGNPDNANVTGEIEPAHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAA QLVTRQLQVSVWHLGTLARRVFLGEVIPLA' WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRITRIQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPGWKHSFVFSG' TFAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWGKVLSSPN LWTDMTLVLH KENKKARNICRIMOSRSRSDGGSEETLPQDH NHHENERRWQQELHREEAYYQFINELNDE DYRLMRDHNLLGTPGETISEELQQRLDGVKE QLASQPDLLRDGTNYRDSEVPRESSHEDSLLE WLNTFRTGNATRSGGNGNOTWRAVSRTNE NNGEFFSLEIHVNHENRGFEIIGEDYTDIPL. DSNRDHTANRQGRSTSSVARRTRSGTSSNFTSN	[	1			1		HNPNAGI.SDI.MSNPVPMPEVOEKFOEOKAK
KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYNMRMLALFVAFAINFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRILPILHTIVISF PCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIJKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT	1	ŀ	1	1	}	1	FEEKEEKEETKSEPEKAEGEDGEKEEKAKED
QQKLLNYFARNFYMRMMLALFVAFANFIELL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQFSEDDIKQQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFIDNSFLYLAWYMT  553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKFVESDADEELLFNIPFTGHVKLK GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKLSRESNYVHSL HISKNFGADTTKVFYGIGRGWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGELEFAHYCFKTHSL EICIKACKNLAYGEEKKKKCMPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHJGTLARRVFLGEVIJPLAT WDFEDSTTOSFRWIPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPGGTLNSFVKGCL3 LPDQQKLRLKSPVLRKOACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH KENKKARNLRNNQSRSRSDGGSEETLPQDH NHHENERRWQOFRLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSQONGNOTWRAVSRTNE NNGEFRFSLEHVNHENRGFEHGEDYTDIPL DSNRDHTARRQQRSTSPVARRTRSSQTSVNFF	1	ſ	1	İ	ŧ		POPONI POI HTHPYGEPPVPFSAFWKKIIAY
FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKEEKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKFKKKDSSLSAVLNSIDVKYQMW KLGVVFIDNSFLYLAWYMT  1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAERVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKPVESDADEELLFNIPFTGHVKLK GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPPDTFSLNRDLTGELEVATKISRFSNVYHLSI HISKNFGADTTKVPYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS 1904 A 4200 1 961 GIPCTEMGNPDNANVTGEIEFAHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQCKRKTGYQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVPLGEVIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGA TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLGTPGEITSBELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNI NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTSSPVARRTRSQTSVNFT		ļ				1	CONTAINS DISTRIBUTED A STEVARADIETT
SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKFPKKDSLSAVLNSIDVKYQMW KLGVVPTDNSFLYLAWYMT    1903	1	1	1	1	i		
FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT  1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAERVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADIDRSKPVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVPYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVUDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA' WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLROSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQORLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGRATRSGONGNQTWRAVSRTNE NNGEFRFSLEIHVNHENGFEHGEDYTDIPL UNDTHTANRQQRSTTSPVARRTRSQTSVNFINE	ļ	ł	1	1	1	1	
TEQPSEDDIKGQWDRL/NTIOSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT  1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGGAERTQQVAA LPLSHGHSHGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVPYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPITQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIPLA' WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL'I LPDQQKLRLKSPYLRQAACPQWKHSFVFSGG' TPAGKRQSLELTVWDQALFGMNDRLLGGT' RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQGELHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTINE NNGEFFFSLEIHVNHENGFEIHGEDYTDIPL WLNTFRRTGNATRSGQNGNQTWRAVSRTINE NNGEFFFSLEIHVNHENGFEIHGEDYTDIPL DDSNRDHTANRQQRSTSSPVARRTRSQTSVNFI	İ	ì			1	ì	
VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT  553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAERVGAARGSAAC AYGLYRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKFVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMELYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKVYVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS HISKNFGADTTKVYVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS EICIKACKNLAYGEEKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVTAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQCERLHREEAYYGFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNCTWRAVSRTNS NNGEFFFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTSSPVARRTRSQTSVNFF	Į.	1	1		1		FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI
VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT  553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAERVGAARGSAAC AYGLYRIDKGRLQCLNESREGSGRGVFKPW ERADDRSKFVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMELYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKVYVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS HISKNFGADTTKVYVIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS EICIKACKNLAYGEEKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVTAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQCERLHREEAYYGFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNCTWRAVSRTNS NNGEFFFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTSSPVARRTRSQTSVNFF	1	1	Ì	1			TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF
FSDAREKKRPKDSSLSAVLNSIDVKYQMW KLGVVFIDNSFLYLAWYMT  1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGCRCAAERVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADIDRSKPVESDADEELLFNIPFITGHVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPIFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIPLAY WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSSV TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHEENERWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNCTWRAVSRTNE NNGEFFFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTSSPVARRTRSQTSVNFF	ł		1		1 .	1	VKRKVMDKYGEFYGRDRISELLGMDKAALD
S53   1903   A   4199   31   767   LPELNGRGGURRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAERVGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERAD\DRSKFVESDADEELLFNIPFTGHFVKLK GIIMGEDDSHPSEMRLYKNIPQMSPDDTER EPDOTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS		ì	1		)	1	FSDAREKKKPKKDSSLSAVLNSIDVKYQMW
1903   A   4199   31   767   LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGCCCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGGVFKPW ERADDRSKFVESDADEELLFNIPFTGHYKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS	-	1	•	1 .	ł		
LPLSHGHSHGGGCRCAAER/VGAARGSAAC AYGLYLRÜKGRLQCLNESREGSGGVFKPW ERADUDRSKFVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPPQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHIFIS  554 1904 A 4200 I 961 GIPCTEMGNFDNANVTGEIEFALHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAGEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPPQQKLRLKSPVLRKQACPQWKHSFVFSGN TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH S555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNTWRAVSRTN NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTNSFVARRTRSQTSVNFI		1002	<del>-  </del>	4100	21	767	
AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERADURSKFVESDADEELLFNIPFTGHVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNYYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 I 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPFFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAY WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAGETDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQCERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNSV NNGEFFFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTNSPVARRTRSQTSVNFF	333	1903	A	4199	31	/6/	I DI SHCHSHGGGGCRCAAER/VGAARGSAAC
ERADIDRSKFVESDADEELLFNIPFTGHVKLK GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554  1904  A 4200  1 961  GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555  1905  A 4211  331  2419  KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNYTWRAVSRTNE NNGEFRFSLEHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRSTNSPVARRTRSQTSVNFF	1	}	1		1		A VCI VI DIDVCDI OCI NESPEGSGRGVEKPW
GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554  1904  A 4200  I 961  GIPCTEMGNFDNANVTGEIEFALHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555  1905  A 4211  331  2419  KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFF		l l	- [	ı	i		AY GLYLKIDKOKLQCLNESKEGSGKGYFKI W
EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 I 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGY TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEHVNHENRGFEHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1	1	1	1	(		
HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS  554 1904 A 4200 I 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPIFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA: WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFGST TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL- DSNRDHTANRQQRST\SPVARRTRSQTSVNFF	1	1	1		1		
CNYEASANPADHRVHQVTPQTHFIS  554  1904  A 4200  1 961  GIPCTEMGNFDNANVTGEIEFALHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA' WDFEDSTTQSFRWHPLRAKADKYEDSVPQSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555  1905  A 4211  331  2419  KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI			-		1		
1904 A 4200 I 961 GIPCTEMGNFDNANVTGBIEFALHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLDSNRDHTANRQQRST\SPVARRTRSQTSVNFT		)	j		ļ		
1904 A 4200 I 961 GIPCTEMGNFDNANVTGBIEFALHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLDSNRDHTANRQQRST\SPVARRTRSQTSVNFT			-		1		CNYEASANPADHRVHQVTPQTHFIS
EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA: WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL- DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	554	1904	A	4200	11	961	GIPCTEMGNFDNANVTGEIEFALHYCFKTHSL
RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA* WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	754	1307	1	1200	1 -	1	
QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555  1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1	1	1	1	1	1	RSSOGKRKTGVORNTVDPTFOETLKYOVAPA
WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555  1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI			1		1 .	1	OLVTROLOVSVWHLGTLARRVFLGEVIIPLAT
NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1		1		1		WDEEDSTTOSERWHPI RAKADKVEDSVPOS
HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI			1		i	1	MOLET LAND YEL ALL DEBLERI UE Y UEGLEUUDGI
LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI			1	İ			NOELI YKAKLYLI SKI KALQEAQEGI DQPSL
TPAQLRQSSLELTVWDQALFGMNDRLLGGTV RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	}	}	J	1	1		HOQLCLVVLGAKNLPVKPDGTLNSFVKGCLT
RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1	1	-		1	1	LPDQQKLKLKSPVLKKQACPQWKHSFVFSGV
LWTDMTLVLH  555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1			)	j	1.	TPAQLRQSSLELTVWDQALFGMNDRLLGGT\
555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1		1			1	RLGSKGDTAVGGDACSQSKLQWQKVLSSPN
555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI		1		1			
MHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	555	1005	- <del> </del> _	4211	331	2419	
DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	222	בטפג ו	A	7211	1331	2717	
QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNI NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI			ì				
WLNTFRRTGNATRSGQNGNQTWRAVSRTNE NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1		ŀ		1		OF YOUDD BUCKIND DESIDE EGREDOLLE
NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPL DSNRDHTANRQQRST\SPVARRTRSQTSVNFI	1		1	Ì	1	}	
DSNRDHTANRQQRST\SPVARRTRSQTSVNFI			ı	1	1		WENTFREIGNAIRSGUNGNQI WKAVSKINP
DSNRDHTANRQQRST\SPVARRTRSQTSVNFI GSSSNIPRTRI ASRGONPAFGSFSTI GRI RNG		1		1	1	Į.	
	1	- 1	- 1	1	1	1	DSNRDHTANRQQRST\SPVARRTRSQTSVNFN
Gabanta Kitanakaya nabata atautaka	}		1	1	1	1	GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						GGAAGIPRANASRTNFSSHTNQSGGSELRQRE
ĺ		[				GQRFGAAHVWENGARSNVTVRNTNQRLEPI
						RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV
				<b>,</b>		QQTTRRSVRRRGRTRVFLEQDRERERRGTAY
1				ł		TPFSNSRLVSRITVEEGEESSRSSTAVRRHPTTT
1	1	1				LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE
	ļ			ļ	,	NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP
1	1					LRRISENELVEPSSVALRSILRQIMTGFGELSSL
						MEADSESELQRINGQHLPDMHSELSNLGTDN
						NRSQHREGSSQDRQAQGDSTEMHGENETTQP
] .	j	ļ	ļ	1		HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH
1		l			1	FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN
					,	SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLGSNIANNG
FF.(	1006	A	4212	3	462	LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR
556	1906	A	4212	3	402	KSPENTEGKDGSKVTKQEPTRRSARLSAKPA
	1	Į	}			PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK
Í		1				QEAGKEGTAPSENGETKAEEIHISRSTVNVST
					1	SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	A	4213	774	507	ARRESCLTLQTSWGHRH\GPPRP\ANFVFLVET
337	1907	^	4213	1 //4	307	GFLHIGQAGHKLPTSGDPPASASQSARITGMS
		ļ	İ			HRTWFLASFLIDSCKNFIVYKIMYTL
558	1908	A	4225	3	1253	TYRHAEREHPETSSATKVSYDYRHKRPKLLD
1336	1503	^	4223	1	1233	GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE
		1				LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC
1		l	Ì	1	l	TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE
		1				SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK
		1				VDVKKTVDTFRVASSYSTERQMSHDLVAVG
	1	1				RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT
		1			Ì	IIHQVKANYFPSPGITLHERFS\KMADIHKADV
		l				NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE
Ì		l	1			QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI
	Į.					ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV
	1					EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ
}		1		ļ	}	KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK
L.		L				KKVP
559	1909	A	4235	1	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL
		1		ĺ	1	LGLSHSPASASQVGGITGTQHHTGLIFGFLIET
1		}			1	EFHHYGQAGLELLTSGDPPALAFQSAGITGVS
	<u> </u>			<u> </u>	<u> </u>	HHAWLQVLNS
560	1910	Α	4246	2	1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ
		1			1	AALVNYSRLSEYAKIEGKKREMYELPVFCLA
1	1		Į			SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR
1	1	1	1	1		LAELVIEVLQQNEEHHAEAFAWWSDLMVEH
		1	1			AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ
						LLINDFLRTGLLICGNGK\FHKHLQDLFAPLVV
	1			1		R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN
Ì	ļ		1			GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK
			[			HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK
		I	1			TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP
			1	1		KL\CSMEMGQEFAKMWHQYHSKIDELIEETV
	1	1		1		KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS
		1	1	1		SFLSFTVKAASKYVDVPKPGMDVADAYVTF
		1		1		VRHSQDVLRDKVNEEMYIERLFDQWYNSSM
			1	1		NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA
			1			SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI
301	1711	^	7231	1300	1 324	FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL
L					<u> </u>	1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

CEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
иепсе		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ŀ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide	·	/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
		· · · ·		1		INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI
		1			į	FMKGNTIKETEAWDFLLAL\GVYPTKKHLIFG
1		1			i	DPKKLITEDFVRQRYLEYRRIPHTDPVDYEFQ
1				ĺ	i	WGPRTNLETSKMKVLKFVAKVHNQDPKDW
		l				PAQYCEALADEENRARPQPSGPAPSS
562	1912	A	4260	1	1498	MYTWLYRFLPTSNMAAKLRSLLPPDLRLQF
	1312	1		_		WLHARLQKCFLSRGCGSYCAGAKASPLPGK
	1	<b>,</b>	1	1		MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ
		1	1	1		WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR
	1	1		l		SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV
1	1				į.	ASGETADVVQTAAEQSFAELGLGSYTPVGLI
				į.	i	ONLLEFMHVDLGLPWWGAIAACTVFARCLIF
				ļ	1	PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA
				1	1	GDHIEYYKASSEMALYQKKHGIKLYKPLILPV
		1			ļ	TOAPIFISFFIALREMANLPVPSLQTGGLWWF
	}	]			1	QDLTVSDPIYILPLAVTATMWAVLELGAETG
	1	1				VQSSDLQWMRNVIRMMPLITLPITMHFPTAV
1.					1	FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR
1					1	VVHDLDKLPPREGFLESFKKGWKNAEMTRQ
					1	LREREQRMRNQLELAARGPLRQTFTHNPLLQ
		1		ŧ		PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG
563	1913	A	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN
1 303	1,,,,	1			ļ	THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK
1				1	1	HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGV\L
1 .			ì			PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF
1	ļ		1	1		LIQENNNTNHTHSHTHTYTETLSFFLYICVNN
	1				i	DRMEWGKSVF
564	1914	A	4270	3	368	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL
1 33.		1				GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF
	}	1		ł	ļ	FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES
		ł		İ		ETYTLSSKMGPDSKPSEGDVFPRTSE
565	1915	A	4288	83	406	RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS
	1				1	PPSALLAPTKPRALGTLRLYECSPELCTTMLP
İ	1	1		ļ	i	PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP
					l l	GQTGASRTPRT
566	1916	A	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL
					1	GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS
1	1	1	-			GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR
1_	1	1	1		1	VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ
1					1	LLKKNGGIVMVTLSMGVLQCNLLANVSTVA
1	1			1		DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E
1				]	1	DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR
			1			VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
	1			1		FHLGASEWTPRLLIWR
567	1917	A	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
1	1	1			1	DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
1	1	1	1			WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ
		1			1	MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
		ĺ			1	VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ
1		j		1		WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
1				1		NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
		1	1			CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD
		1	-			INEAYVETLKHCFMMPQSLGVIGGKPNSAHY
				.	1	FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES
Į.	!				1	FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF
	1				1	GAECCLGMTRKTFGFLRFFFSMLG
568	1918	HA-	4300	2012	1843	SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS
1 2 3 3	1	1	1	1		LMSVLIPKLPQLHGVRIFGINKY
569	1919	A	4302	186	531	WTFCLFL/WWVPESARWLLTQGHVKEAHRY

				20.00	T 10 10 10 10 10 10 10 10 10 10 10 10 10	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	i	USSN		to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	ŀ	09/496	correspondi	acid residue	O=Glutamine, R=Arginine, S=Serine,
иепсе		ł	914	ng to first		T=Threonine, V=Valine, W=Tryptophan,
		(	[	amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł		1	residue of	sequence	/=possible nucleotide deletion, \=possible
	ł		l	peptide	1	nucleotide insertion
				sequence		nucleoude insertion
	)	1	1	Ì	Į.	LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI
		•		ĺ		CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT
		l	<u> </u>			VTESKLEAEGKTKEKAREKERKKKS
570	1920	A	4308	3	869	RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS
				1	ł	GKRNKLRVYYLSWLRNKILHNDPEVEKKQG
		1		1		WTTVGDMEGCGHYRVVKYERIKFLVIALKSS
	i	İ	l		1	VEVYAWAPKPYHKFMAFKSFADLPHRPLLV
	ļ	1	1	1	1	DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY
	1	1	1		1	DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE
	1	}	3	ļ	}	DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC
	Į.	[	1	į.	Į.	SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA
				1		QRLKFLCERNDKVFFASVRSGGSSQVYFMTL
	1	-		1		NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE
] -,.	1			í		KEHREFRAKTNRDLEIKDQEIEKLRIELDESK
1				1		QHLEQEQQKAALAREECLRLTELLGESEHQL
		1		İ		HLTROEKDSIQQSFSKEAKAQALQAQQREQE
1		1	,	1	}	LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT
		1	1	1		KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
312	1722	1	1310	1.		DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
Ì	1	1	1	1	1	WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ
ł	1	1	1	}		MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
1	1			[		VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ
1	1	1	1		1	WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
	1	į.	ł	ļ		NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
			,		1	CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T
	i	1	ł			DINEAYVETL\KHCFHGWPQFPG/VVHREGK
						PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG
	İ	Į	1	1		CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH
	<b>\</b>	ĺ	İ	1		LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
L	<del> </del>		4222	1 262	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD
573	1923	A	4333	363	1000	MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS
	i				i	LFLTIPNLAISWEGHIVVYSSTEEKTTLK\ERM
Ì					ŀ	HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG
1			İ			SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT
}	J	}	• }	]		KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN
		1			<b>\</b>	FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD
1		]				1 2 2 7 7 1
					<del>  1024                                   </del>	FSKGSK
574	1924	Α	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL
1		1	l	1	į.	LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK
ì	İ		1			LPSARAKIRITSSPIFITFYILVFVVALVGIARA
	1	1			<b>\</b>	VVSMTVSTSNAATVADKILWEITRFFLLAIEL
-	1	1	ļ.			SVIILGLAFGHLESKSSIKRVLAITTVLSLAYSV
I		1	j	1	]	TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL
1		-	ļ	İ	1	VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY
1	1		-		-	VYAGILALLNLLQGLGSVLLCFDIEGLCCVD
1						ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	A	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV
1 - 1 -		1				SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM
1		I		1		KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT
1		- 1	-	1	1	VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP
1	- 1	l	i	1		ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP
	1	i	- [	)	1	ODODKLSKEDVLSFIQMHRA
		3				
576	1004	<del>-   , -</del>	1365	60	500	OVEGROGREVKRTAWRISPVWRPARCKRKS1
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST POP/PE/PGAOOOERHROGEAPMOALDPRAEP
576	1926	A	4365	69	500	PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP
576	1926	Ā	4365	69	500	PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ
576	1926	Ā	4365	69	500	PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP

	0000	Mat	CEA	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid,
nuci-	peptide	nou	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
otide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
jence	ucirco	Ì	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
LCHCC	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	į	· .		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			}	peptide	1 •	/=possible nucleotide deletion, \=possible
				sequence	<u> </u>	nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
	1					ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
						FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
		1				SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
		1	1			SGWSRTPDLR
579	1929	A	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
		1	}		ì	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
	1		1	i	\	CWPGWSSTPDLK
580	1930	A	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
		1				\VFKKGNHILHELFQNKEBGAFPNS/FYEASFT
	İ	1		1		LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
				<u> </u>		QLKSSDL
581	1931	Α	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
		i		1	1	RLPHATTVDHRPQHRWLETCNAPPQLIQGKA RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
		}	.			RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
				1	}	LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
		ì	1			VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
		}	İ	1	-	DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
	1 .	1	ľ	1		SPE SPE
		<del> </del>	1404	104	449	VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
582	1932	Α	4424	194	447	LEQELLEHGRDAASVQAATSVQAMQGKTTL
	1	İ		1		PS\OGPLORPSRLVFT\DVANAIHV
***	1022	<del>                                     </del>	4435	1 .	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
583	1933	Α	4433	1	100	PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
				1		SAPPALLODTSV
584	1934	A	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
304	1934	Α.	1777	1	""	APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
		1	1			APATQHSQAGPATGQAYGPHTYTEPAKPKK
		1				GOOLWNRMKPAPGT\EV\$SST\$R\$DPLLLPPR
		ı	İ	1		ALAPTORASTVVLAPSPT/SEKVQNHSGSSAR
		1		}		GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
		1		}	į	QGPHGKAAQGGAAGAAGRLGLYH
585	1935	A	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
	1	1		1	<u> </u>	SIFDDFAHYEKRQ
586	1936	A	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
		1		,	1 -	FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
		Ţ	1		-	INLEQCEEIEALSMAFYSSPEILRVPDSRKKVP
	1	- [	-	l		TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
	İ	1	1		ì	FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
		1			1	LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
		- (		i	1	PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQL'
				1		ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
1		İ		1		LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
	1	1		1		PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
1		1	1			LPGALVIEVKWTKYGSLLNPQAKIVNVTANI
	1	1	f	1		SSSFPEANSGNERTILISTAVTFVDVSAPAEAC
						FRAPPAINARLPFNFFFPFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
				İ		CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCI
1						NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPI
1				İ		CPANFCIII/DFLVETGFHHVGQASHELLTSGI
		1				PPTSASQSAGITGMSYHTWFGES
589	1939	A	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
1	1	- 1	1			PPVELPWAPRRGHRLSPADDELYQRTRISLLO
i	l l					
						REAAQAMYIDSYNSRGFMINGNRVLGPCALI PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	residue of	sequence	/=possible nucleotide deletion, \=possible
		ļ		peptide	1	nucleotide insertion
		<u> </u>	ļ	sequence	ļ	VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ
	l	1		]		DTPNACATFNFLCHEGRVTGAALIPPPGGTSL
				1		TSLGOAAO
		<u> </u>	1400	ļ <u>.                                    </u>	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
590	1940	A	4492	1	4/2	PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
	]	]	1	J	1	RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR
			1		İ	ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
		ļ	1		ļ	VPHQGGLPGPIRVAPSSAGQREASQGPPGR
	1044	<del> </del>	4495	1444	1116	IAARFTLAKTWNQLKRP\TMIDSIKKTR\YIYT
591	1941	Α	4493	1444	1110	MEYYADTERNEIMSF\AGTWVELEAIILSKLM
			ļ	}		LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL
	ł	1		1	1	EPWDSSCFPHPSSGV
602	1942	A	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
592	1742	1^	4470	12	1 ***	VCGGLSLLANAWGILSVGAKQKKWKPLEFL
		i	İ	1	1	LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
	i	1	1	1		EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
	1	ļ				MWMVCWPVNYRLSNAKKQAGHTVMGIWM
		İ			1	GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
-		١.	ì			GLGFGVCFLLLVGGSVAMGVICTAIALFQTL
		1	1	Ì		AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI
	İ	1		1		DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL
		ļ		i		GPFSLADTHLSDLPYTWGDRDSGGACVM
593	1943	TA.	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC
	1	1			ł	HFPASASQVAGTTHARHHTQLIF\AFLVENGL
-				1		C
594	1944	A	4507	1327	647	KMAGGVRPLRGLRALCRVLLFLSQFCILSGG
					1	ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS
	İ	1	-			CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFU
	1		1	-		NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ
	İ	İ	ļ			RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
		-	1			GGYKWVYGLWLLRHHPRWGLGADRF\YLGP
						VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG
						PADGSLYI FFFKMESYSVARLECSGAISAPCNLHLLGSNN
595	1945	Α	4512	533	264	SPASASRV/AGNIGARHHTQQIFVLLVQMRVH
		1		1		YVGQDGLDLL/NLMIHPPRSPKVLGLQA
					1	HASDHLYPNFLVNELILKQKQRFEEKRFKLD
596	1946	A	4513	3	1674	HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
	1			1		NLMLELLVQKKKQLEAESHAAQLQILMEFLK
ļ		Ì		l l		VARRNKREQLEQIQKELSVLEEDIKRVEEMS
i	1			1	1	GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP
}		-				PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
Į.			1	1	}	DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS
1		1		- [	1	KF/TRYNSVRPL\ATLSYASDLYNGSQYKSLV
1		1		1		FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA
1		i		i		VDIHYPENEMTCNSKISCISWSSYHKNLLASS
1		Ì				DYEGTVILWDGFTGQRSKVYQEHEKRCWSV
			1			DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
	1	1	ł	1		ASIEAKANVCCVKFSPSSRYHLAFGCADHCV
i		1	1	1	1	HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG
1	i	- 1				EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN
1	1			1		EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
1		1		1		KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV
1		}		}	1	CWRALPDGESNVLIAANS\QGTI\KVLELV
I.	1			536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS
	<del>-  </del> -				1.044	
597	1947	A	4518	130	1	A SOVAEITS VRHHTWLIFCILGOMGFHHVGE
597	1947	A	4518	330		ASQVAEITSVRHHTWLIFCILGQMGFHHVGE
597	1947	A	4518	1	384	ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF

NO. of   No. of   N		- And		OCO.	Dandista J	Dundistad and	Amino acid sequence (A=Alanine C=Cysteine,
in mucleotide could sequence    Peptide could sequence							D=Aspartic Acid. F=Glutamic Acid.
Seq-   URSN   Osation   Corresponding   Seq-   URSN   Osation   Seq-   URSN   Osation   Osatio			noa				F=Phenylalanine, G=Glycine, H=Histidine.
uence uence uence			1				
			l				
mino acid residue of peptide sequence	-	uence	1				
Prisidue of peptide   Sequence	uence		1	914			
		i	]				
		1		1		sequence	
SQL   1949   A   4526   366   776   MGQRAYASHHGYITYSGFLVYPDLAPAA  ASELL			ļ				
1949   A   4526   366   776   MGQPAYSHGKYTTFSGFLVYPDLAPAA   ASELL				<u> </u>	sequence		
ASELL						1	KKKEMQSQSVMLALKKGDAV WLLSHDIDG
1949   A   4526   366   776   MGQPAPYAEGPIQGGDAGELCKCDPI. NPEAVCEAGPTAMPGIANGMESCS VQWRDPGSLHPPPLGFKRFSCLSLPSS HAPPHPANFCIPSRQVSPCWPGWSCR. PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFKVLGLQA   PPPWLFSRSYVAQAGVGWL   LISTGVLTGSVSDLGRTPP   RANFFYPVLFSRSYVAQAGVGWL   LLISTGVLTGSVSDLGRTPP   PRVQESHFLESEPSRGISDNYTLALIT   VGSRKARFALIMILTWRAEQEGGMQI   SKLSDSWQPRSLDIEVAAYALLSHFIQ   GPPMRWLSQRNSLGGFASTQDITV.   EFAALMNTERTNIQVTVTGPSSPSP   HNRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAMSGV/SISAN   CQLNVYNYKASGSSRRRSIGNSLGAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLLQTAELADGTAAVALLSHFIV   PHRLQTAELADGTAA		1	i	ì		ļ.	1
NPEA/VCEAGTPAMPQITAWRQMESS		l	<u> </u>				
VQWRDPGSLHPPIGFKRRSCLSIPS	599	1949	Α	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDrLvri3r
HAPPIPANFCIFSRDQVSPCWPGWSR.			1			1	NPEAVCEAGIPAMPQIAWRQMESCSIAQAG
PPPWLPKVLGLQA		}	1				
600		1	ł	}	1		
DPPASASRVAGTTGARHHTPULTVETV   WILARDGURLITISSPPASASQSSWD   RLANFFVFLVETGSRYVAQAGVQWL   LLISTGVLTCSVSDLGRFTPP   HEVQESBIFLESEPESRGISDNYTLALIT   VGSPKAKEALNMLTWRAEQEGMQ   SKLSDSWQPRSLDIEVAAYALLSHFLL   GPIMR WLSRQRNSLGGPASTQDITV.   EFAALMMTERTNIQVTVTOPSSPPVH   HNILLIQTAELADGTANGSVJSISAN   CQLNVYNVKASGSSRRRSIGNQE   AVKENKDDLNHVDLNVCTSFSGPGR   MEVNILLSGRWPSEAISLSETVKKVE   LNLYLDSVNETQFCVNIPAVRKVS   SVSIVDYYFPRRQAVRSYNSEVKLSS   VQRLPSL   WRAFGRPALRPLPLPFLLLLLSSPWC   SGSILFKPANITFLSINMKNNLQWTP   VKVTYTVQYF1YQGKKWLNKSECRN   DLSAETSDVEHQYYAKVKAHWGTKC   SGRFYPFLETQIGPPEVALITDEKSIS   EKWRNPEDLPVSNQQIYSNLKTNV   KSNRTVSQCVTNHTUVLTWLEENTI   ESFVPGPPRRAQPSEKQCARTILOGS   IIFWYVLPISITVFLFSVMGYSIYRYIH   HPANLILIYGNEFDKRFFVPAIEXIV   NISDDSKISHQDMSLLGKSSDVSLN   NLRPPGEEEVKHLGYASHLMEIFCC   EGTSFTQGESLSRTIPPDKTVUEYYD   CAGPEEQELSLCEVSTQGTLLESQA   GPQTLQYSYTPQLQDLDPLAQEHTD   EPSTTLVDWDPQTGRLCIPSLSSPQI   PSGDGGGGEGELSKTEPPDKTVEYYD   CAGPEEQELSLCEVSTQGTLLESQA   GPQTLQYSYTPQLQDLDPLAQEHTD   EPSTTLVDWDPQTGRLCIPSLSSPQI   PSGDGGGGEGGGGGGGLSKEPPAPDRPP   LMQPMEEWGLYVQMEN   GMIAPYYEDSDLKDLSKSVLQSPVS   LQAVIAGDLMKLIESYKNGGSLIQG   LHYAAETGNGEIVKYILDHOPSELLL   TGETALHKAACQRNRAVCQLLVDA   TDSKGKTPQERAQAGDPDAAAYT   KVIGHEDLETAV   KVIGHEDLETAV   KVIGHEDLETAV   KVIGHEDLETAV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIFPEPPAPBARA   QDNKVQNGSLHQKDTVHDNDFEPY   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIFPEPPAPBARA   QDNKVQNSSYFSNSDPVLSSYYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIFPEPPAPBARA   QDNKVQNSSAYSSYYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIFPEPPAPBARA   QDQQTQSSAYGSSYTYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIFPEPPAPBARA   QQQQTQSSAYGSSYYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRRIFPEPPAPBARA   QQQQTQSSAYGSSYYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRRIFPEPPAPBARA   QQQQTQSSAYGSSYYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRRIFPEPPAPBARA   QQQQQTQSSAYGSSYYPPSSLGGTV   STAGDPPPYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIPPEPP   STAGDPPTYLTTYGQLSNGDHHFMI   QPGGLGNNIVQHRNIPPEPP   STAGD		l	l				PPPWLPKVLGLQA
WILARDGLKILTSSDPPASASQSSWD   RLANFPVLVETGSKYVAQAGVQWL   LLISTGVLTCSVSDLGRFTPP	600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS
RLANFFYPLETGRSYVAGAGVQWLL LLISTGVLTCSVSDLGRFTPP  601 1951 A 4533 1460 403 HEVQESIHFLESEFSRGISDNYTLALIT VGSPKAREALMMLTWRAGPGGMQ SKLSDSWQPRSLDEWAYALLSHFLG GPIMRWLSRQRNSLGGFASTQDTTV. EFAALMNTERTNIQVTVTGPSSPSPYW HNRLLQTAELADGTANGSVSISAM CQLNVYYNKASGSSRRRSIGNQEA AVKENKDDLMHVDLNVCTSSGPGR MEVNLLSGFMYPSEAISLSETVKKVE LNLYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPPPLLILLLSSPW VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQYFIYGQKKWLNSECRN DLSAGTSDYEHGYYAKVKAMGTKC SGRPYPLETQIGPEVALTDEKSIS EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFYPGPPRAQPSEKQCARTLKDQS IIFWYVLPSITVTSFSVMGYSIYRYH HPANLLIJGNEFDKRFFVPAEKIV NISDDSKISHQDMSLLGKSSDVSLN NLRPPQEEEVKHLGYASHLMEIFCL EGTSFTQQESLSRTIPPDKTVEYSYD CAGPBEQLSLGEVSTGGTLLESQA GPQTLQYSYTPQLQDLDAQEHTD EPSTTLVDWDPGTGRLCPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  603 .1953 A 4543 3 600 YSAVEFVEQASGDWNPPALRKR GMMAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLESYKNGGSLIQC LHYAABTGNGEIVKYLDHGPSELLL TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQAGDPDLAAVT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STAGDPPIPYLTTYQQLSNGDHHFMI QPGGLGNNIYQHRPNFFPDNAFSAV GSNSYSPSMSDPYLSSYYPPSIGPFYSI STAGDPPIPYLTTYQQLSNGDHHFMI QPGGLGNNIYQHRPNFFPPNAFSAV GGGGGGNSTAYSTYTPPSSLGGTV			1	ļ.		]	DPPASASRVAGTTGARHHTQLIFVFLVETGFH
LLISTGVLTCSVSDLGRTPP	İ	ł	1	1		}	MLARDGLKLLTSSDPPASASQSSWDYRREPP
601 1951 A 4533 1460 403 HEVQESIHFLESEFSRGISDNYTLALTI VGSPKAKRALMMLTWRAEQEGGMQI SKLSDSWQPRSLDIEVAAYALLSHFLG GIPIMRWLSRQRNSLGGFASTQDITV. EFAALMNTERTNIQVTVTGPSSPSPVW HNRLLLQTAELADGTANGSVSISAM CQLNVYNVKASGSSRRRSIQNG2A AVKENKDDLMHVDLNVCTSESGFGR MEVNLLSGFMVPSEAISLSETVKKVE LNLYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYFERRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPPLPFLLLLLSSPWC SGGJKPKANITHLSINMKNVLQWTP VKVTYTVQYFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPFLETQIGPPEVALITIDEKSISS EKWKRNPEDLPVSMQQIYSNLKYNW KSNRTWSQCVTNHTVLTWLEPNTT ESFYPGPPRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGSYINYTH HPYANLLLIYGNEFDKRFFVPAEKIV NISDDSKISHQDMSLLGKSSDVSSLM NLRPPQEEEEVKHLGYASHLMEIFCL EGTSSTQQELSICJEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSTTLVDWDPQTGRLCIPSLSSFDQI PSECDGLGEEGLLSRLYEEPAPDRPPLMQFMEEWILYYQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRM GMIAPYYEDSDLKDLSHSRVLQSPWS LQAVIAGDLMKLESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLT TGETALHKAACQRNRAVQLLVDA TDSKGKTPQERAQQAIGDPDLAAYTYKVIGHEDLETAV QLSVGSTPSMSDPYLSSYYPPSIGPFYSS STAGPPPPYLTTYQQLSNDGHHFMI QPGGLGNNIYQHFFNDAFFAX QSNSYPSMSDPYLSSYYPPSIGPFYSS STAGPPPPYLTTYQQLSNDGHHFMI QPGGLGNNIYQHFFNDAFFAX QGQQTQSSAYGSSYTYYPPSSLGGTV		1	1	İ			
VGSPKAKRAINMLTWRAEDEGGMOJ SKLSDSWOPRSLDEVAAYALLSHFLG GIPIMRWLSRQRNSLGGFASTQDITV. EFAALMNITERTNIQVIVTGPSSPSPYM HNRLLLQTAELADGTANGSV/SISAN CQLNVVYNVKASGSSRRRSIQNQEA AVKENKDDLNHVDLNVCTSPSGPGR MEVNLLSGFMVSPEAISLSETVKKVE LNLVLDSVMETQFCVNIPAVRNFKVS SVSIVDYYEPRRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPPLILLLSSPW VSGGLPKPANITFLSINMKNNLQWTP VKVTYTVQYFTYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPLETQIGPPEVALTIDEKSIS EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNIT ESFVYGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYHH HPANLLLIYGNEFDKFFVPAIEKIV NISDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEVKHLGYASKILMEFCL EGTSFTQGELSLRTPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAGEHTD EPSTITLDWDPDGTGRLCIPSLSSFDQI PSEGDGLGEGGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMAPYYEDSDLKDLSHSRVLQSPWS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYLDHGPSELL TGETALHKAACQRNRAVQLVDAA TDSKGKTPQERAQQAIGDPLAAYT YKVIGHEDLETAV QSNSYPSMSDPYLSSYYPPSIGPFYSI STAGDPPIPYLTTYQGLSNGDHIFMI QPGGLGNNIYQHRFNFPNAFSAV QSNSYPSMSDPYLSSYYPPSIGFPYSI STAGDPPIPYLTTYQGLSNGDHIFMI QPGGLGNNIYQHRFNFPNAFSAV QGQGTQSSAYSSYSTYPPSSLGGTV	ľ	i	1		·	1	
SKLSDSWOPRSLDIEVAAYALLSHFILL GIPIMRWLSRQRNSLGGFASTQDITV. EFAALMNTERTNIQVTVTGPSSPSPVH HNRLLLQTABLADGTANGSVISIASM CQLNVYNVKASGSSRRRRSIQNQEA AVKENKDDLNHVDLNVCTSFSGPGR MEVNLLSGFMVPSEALSLETVKKVO LILVLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPPILLILLSSPW VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQFFIYGKKWLNKSECRN DLSAETSDVEHQYVAKVKAIWGTKC SGRFYPFLETQIGPPEVALTTDEKSISS EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFYPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HPANLLIYGINEFDKRFFVPAEKIVV NISDDSKISHQDMSLLGKSSDVSSIN NLRPPQEEEVKHLGYASFILMEFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLODLDPLAQSHTD EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLVEEVAPDRPP- LMQPMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWPALRKRN GMAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGSSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALKRAACQRNRAVCQLLVDA TDSKGKTPGERAQA/GDPDLAAVT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STRAGDPIPYLTTYGQLSNGDHIFMI QPGGLGNNIYQHRFNFFPENPAFSAN QGQCTQSSAYGSSYTYPPSSLGGTV	601	1951	A	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS
GIPIMRWI.SRQRNSLGGFASTQDITV.  EFAALMNTERTNIQVTVIGPSSPSPVW. HNRLLLQTAELADGTANGSV/SISANG CQLNVVYNVKASGSSRRRSIGNQEP.  AVKENKDDLNHVDLNVCTSFSGPGR MEWNLLSGFWYPSEAISLSETVKKVE LNLYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPLPPLLLLLSSPWG VSGGLPKPANITT.SINMKNVLQWTP VKVTYTVQYFTYGQKKWLNKSECRN DLSAETSDVEHQYVAKVKAIWGTKC SGRFYPFLETQIGPPEVALTTDEKSISS EKWKRNEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEENTI ESFVPGPPRRAQPSEKQCARTLKDQS IIIFWYVLPISITVFLESVMGYSIYRYH HPAANLLLYGNEFDKRFEVPALEKIVV NISIDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCL EGTSFTQQESLSRIPPPDKTVEYSYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGGGLGEGLLSRLYEEPAPDRPP- LMQPMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYLDHGPSELLC LHYAAETGNGEIVKYLDHGPSELLI TGETALHKAACQRNRAVCQLLVDAM TDSKGKTPQERAQQAGDPDLAAVTT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDMKVQNGSLHQKDTVPDNDFEPY QSNSYPSMSDPYLSSYYPPSIGGFTYSL STAGDPPPLTTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAN QGQCTQSSAYGSSYTYPPSSLGGTV	1	1	1		Į.	1	VGSPKAKEALNMLTWRAEQEGGMQFWVSSE
EFAALMTERTNIQVTVTGPSSPSPYK   HNRLLIQTAELADGTANGSV/SISAN( CQLNVYNVKASGSSRRRSIQNOEA AVKENKDDLNHVDLNVCTSFSGPGR MEVNLLSGFMVPSEAISLSETVKKVE LNLYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRQAVRSYNSEVKLSS VQRLPSL   MRAPGRPALRPLPPLLLLLSSPWC VSGGLFKPANITFLSINMKNVLQWTP VKVTTVQFTQGKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPFLETQIGPPEVALITDEKSISS EK WKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISTYTFLFSVMGYSTYRYH HPVANLILIYGNEFDKRFFVPAIEKIVN NISUDSKISHQDMSLLGKSSDVSAIL NOES IIFWYVLPISTYFLFSVMGYSTYRYH HPVANLILIYGNEFDKRFFVPAIEKIVN NISUDSKISHQDMSLLGKSSDVSAIL OG GCTLQYSYTPQLQDLDFLAQEHTD EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEGLLSRLYEEPADPRPP LMQFMEEWGLYVQMEN    603	1	1	1		]		SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE
HNRLLLQTAELADGTANGSV/SISANG CQLNVVYNVKASGSSRRRRSIQNQEA AVKENKDDLMHVDLNVCTSFSGPGR MEVNLLSGFMVPSEAISLSETVKKVE LNLYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPLPPLLLLLSSPWC VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQYFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRPYPFLETQIGPPEVALITDEKSIS' EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEFNTI ESSPVPGPPRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HPANLLLIYGNIEPDKRFTVVAAKKIV NISDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCL EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEGELSLGEVSTQGTILESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSTTLVDWDPQTGRLCIPSLSSFDQI PSECDGLGEEGLLSRLYEEPAPDRPP- LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRM GMIAPYYEDDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLL TGETALHKAACQRNRAVQLLVDA TDSKCKTPQERAQQAIGDPDLAAVT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAN QCGQTGSSAYGSSYTYPPSSLGGTVY							GIPIMRWLSRQRNSLGGFASTQDTTVALKALS
CQLNVYNYKASGSSRRRSIQNQEA AVKENKDDLNHVDLNVCTSFSGPGR MEVNLLSGFMVPSEAISLSETVKKVE LNLYLDSVNETOFCVNIPAVRNFKVS SVSIVDYYEPRRQAVRSYNSEVKLSSI VQRLPSI  602 1952 A 4540 1963 295 MRAPGRPALRPLPPPLLLLLSSPWC VSGGLPKPANITFLSIMKNVLQWTP VKVTYTVGYIYGOKKWLKKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRPYPFLETQIGPPEVALITIDEKSISN EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESSPVGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HPANLLLIYGNEFDKRFVPAAEKIV NISIDDSKISHQDMSLLGKSSDVSLN NLRPPQEEEEVKHLGYASHLMEIFCL EGTSFTQCPSLSRTIPPDKTVUEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLL TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQAIGDPDLAAVT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN	1	1:	1	1			EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT
AVKENKDDLNHVDLNVCTSFSGPGR MEVNLLSGFMVPSEAISLSETVKKVE LNLYLDSVNETQFCVMIPAVRNFKVS SVSIVDYYPPRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRALRPLPLPPLLLLLSSPWC VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQYFTYQKFWLKKSECRN DLSAETSDYBHQYYAKVKAIWGTKC SGRFYPFLETQIGPPEVALTIDEKSISN EKWKRNEDLPYSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIIFWYLPISTIVTFLFSVMGYSTYRYHH HPANLLLIYGNBFDKRFFVPAEKIV NISUDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFGL EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQLSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSTILVDWDPQTGRLCIPSLSSFDQI PSGGDGLGEEGLLSRLYEEPAPDRPP- LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALKRN GMMPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYLLDHGPSELLD TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQA\GDPDLAAVT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVONGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPIPYLTTYGQLSNGDHFMI QPGGLGNNIYQHRNFFPENPAFSAN QPGGTGNSAYGSSYTYPPSSLGGTV		1	1		1		HNRLLLQTAELADGTANGSV/SISANGFGFAI
MEVNLLSGFMYPSEAISLSETVKKVE LNLYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRRQAVRSYNSEVKLSS VQRLPSL  602 1952 A 4540 1963 295 MRAPGRPALRPLPPLLILLLSSPWC VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQYFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRYPFFETTQIGPPEVALITIDEKSIS' EKWKRNEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFYPGPPRRAQPSEKQCATIKDQS IIFWYLPISITVFLFSVMGYSIYRYH HP\ANLILIYGNEFDKRFFVPA\EKIV\ NISDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEVKHLGYASHLMEIFCL EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDV EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP- LMQFMEEWGLYVQMEN 603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRM GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESVKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLC TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQA\GDPDLAAVTI YKVIGHEDLETAV 604 1954 A 4548 3 938 QDNKYVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSI STAGDPPIPYLITTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAN QPGQCTQSSAYGSSYTYPPSSLGGTV				ļ	Ì		CQLNVVYNVKASGSSRRRRSIQNQEAFDLDV
LINYLDSVNETQFCVNIPAVRNFKVS SVSIVDYYEPRQAVRSYNSEVKLSSI VQRLPSP		1	İ	ļ	1		AVKENKDDLNHVDLNVCTSFSGPGRSGMAL
SVSIVDYYEPRRQAVRSYNSEVKLSS   VQRLPSI	1		1				MEVNLLSGFMVPSEAISLSETVKKVEYDHGK
602 1952 A 4540 1963 295 MRAPGRPALPLPLPLLLLLSSPWC VSGGLPKPANITFLSINMKNVLQWTP VSTYTTVQYFITYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPFLETGIGPPEVALTTDEKSISE EK WKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEFNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYNL HPYANLLITYGNEFDKRFFVPALEKIV NIS\DDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEVKHLGYASHLMEIFCL EGTSFTQQESLSRTIPPDKTVIEFYEYD CAGPEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSITLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYLLDHGPSELLD TGETALHKAACQRNRAVCQLLVDAYTOSKGKTPQERAQQAIGDPDLAAYTI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQTQSSAYGSSYTYPPSSLGGTV	1			1			LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA
602 1952 A 4540 1963 295 MRAPGRPALRPLPEPLLILLSSPWG VSGGLPKPANITHLSINMKNVLQWTP VSGGLPKPANITHLSINMKNVLQWTP VKVTYTVQYFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPFLETQIGPPEVALTTDEKSISY EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HPVANLILIYGNEFDKRFFVPALEKIV NISUDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTGGTLLESQA GPQTLQYSYTFQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEQLLSRLYEEPAPDRPPI LMQFMEEWGLYVQMEN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLIQG LHYAAETGNGEIVKYILDHGPSELLG TGETALHKAACQRNRAVCQLLVDAYT TDSKGKTPQERAQQAIGDPDLAAYTT YKVIGHEDLETAV QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPIPIYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QCQQTQSSAYGSSYTYPPSSLGGTVY	-	Į.	1	1		Ì	SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD
VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKG SGRFYPFLETQIGPPEVALITIDEKSIS\S EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYTH HPANLILIYGNEFDKRFFVPAIEKIV\ NISUDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD\ EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDCLGEEGLLSRLYEEPAPDRPP\ LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDA\ TDSKGKTPQERAQQA\GDPDLAA\YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPIPIVLTTYGQLSNGDHIFMI QPGGLGNNIYQHRFNFFPENPAFSAV QCQQTQSSAYGSSYTYPPSSLGGTV			i	ì	1		VQRLPSL
VSGGLPKPANITFLSINMKNVLQWTP VKVTYTVQFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKG SGRFYPFLETQIGPPEVALITIDEKSIS\S EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYTH HPANLILIYGNEFDKRFFVPAIEKIV\ NISUDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD\ EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDCLGEEGLLSRLYEEPAPDRPP\ LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDA\ TDSKGKTPQERAQQA\GDPDLAA\YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPIPIVLTTYGQLSNGDHIFMI QPGGLGNNIYQHRFNFFPENPAFSAV QCQQTQSSAYGSSYTYPPSSLGGTV	602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLLSSPWGRAVPC
VKVTYTVQYFIYGQKKWLNKSECRN DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPFLETQIGPPEVALTTDEKSIS' EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYH HP'ANLILIYGNBEFDKRFFVPAIKIV\ NISDDSKISHQDMSLIGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSITLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP- LMQFMEEWGLYVQMEN LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRM GMIAPYYEDSDLKDLSHSRVLQSFVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQNRAVCQLLVDAY TDSKGKTPQERAQQA\GDPDLAAYT YKVIGHEDLETAV  604  1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	002	1,752	1 ' '				VSGGLPKPANITFLSINMKNVLQWTPPEGLQG
DLSAETSDYEHQYYAKVKAIWGTKC SGRFYPFLETQIGPPEVALTTDEKSIS' EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HP\ANLILIYGNEFDKRFFVPA\EKIV\ NISIDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPPI LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDAI TDSKGKTPQERAQQA\GDPDLAA/YI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	İ	1	1	1	}	1	VKVTYTVQYFIYGQKKWLNKSECRNINRTYC
SGRFYPFLETQIGPPEVALTTDEKSISY EK WKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTW\LEPNTI ESFVPGPPRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HP\ANLILIYG\NEFDKRFFVPA\EKIV\ NIS\DDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDLGEEGLLSRLYEEPAPDRPPI LMQFMEEWGLYVQMEN  4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLIQ TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQA\GDPDLAAYTI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHIFFM QPGGLGNNIYQHRFNFFPENPAFSAN QGQQTQSSAYGSSYTYPPSSLGGTV		1	1	1	ļ		DLSAETSDYEHQYYAKVKAIWGTKCSKWAE
EKWKRNPEDLPVSMQQIYSNLKYNV KSNRTWSQCVTNHTLVLTWLEPNTI ESFVPGPPRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HPVANLLIYGINEFDKRFFVPAIEKIV NISIDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPPI LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWPALRKRM GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLE TGETALHKAACQRNRAVCQLLVDAI TDSKGKTPQERAQQAIGDPDLAA/YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN QPGGLGNNIYQHRFNFFPENPAFSAN			1				SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP
KSNRTWSQCVTNHTLVLTW\LEPNTI ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HP\ANLILIYG\\NEFDKRFFVPA\\EKIV\ NISUDDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD\ EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP\ LMQFMEEWGLYVQMEN  603		1		ŀ	1		EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT
ESFVPGPPRRAQPSEKQCARTLKDQS IIFWYVLPISITVFLFSVMGYSIYRYIH HP\ANLILIYGINEFDKRFFVPA\EKIV\ NIS\DSKISHQDMSLLGKSSDVSSLN NLRPPQEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP\ LMQFMEEWGLYVQMEN  4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDA\ TDSKGKTPQERAQQA\GDPDLAAYTI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV		ĺ	1	1	l	<b>,</b>	KSNRTWSQCVTNHTLVLTW\LEPNTLYCVHV
IIFWYVLPISITVFLFSVMGYSIYRYIH HP\ANLILIYG\NEFDKRFFVPA\EKIV\ NIS\DDSKISHQDMSLLGKSSDVSSL\ NRRPPQEEEVKHLGYASHLMEIFCD EGTSFTQQESLSRTIPPDKTVIEYEYD\ CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD\ EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP\ LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKR\ GMIAPYYEDSDLKDLSHSRVLQSPVS\ LQAVIAGDLMKLIESYKNGGSLLIQG\ LHYAAETGNGEIVKYILDHGPSELLD\ TGETALHKAACQRNRAVCQLLVDA\ TDSKGKTPQERAQQA\GDPDLAAYTI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY\ STAGDPPIPYLTTYGQLSNGDHHFMI\ QPGGLGNNIYQHRFNFFPENPAFSAV\ QGQQTQSSAYGSSYTYPPSSLGGTV\		1		1	İ		ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK
HP\ANLILIYG\NEFDKRFFVPA\EKIV\ NIS\DDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEEVKHLGYASHLMEIFCD EGTSFTQESLSRTIPPDKTVIEYYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPPI LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDAG TDSKGKTPQERAQQA\GDPDLAAYT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV			1	Ì			IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK
NIS\DDSKISHQDMSLLGKSSDVSSLN NLRPPQEEEVKHLGYASHLMEIFCE EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTD: EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP- LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLE TGETALHKAACQRNRAVCQLLVDA- TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHFFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV			1			1	HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL
NLRPPQEEEEVKHLGYASHLMEIFCE EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDDS PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLE TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV							NIS/DDSKISHQDMSLLGKSSDVSSLNDPQPSG
EGTSFTQQESLSRTIPPDKTVIEYEYD CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  7SAVEFVEQASGISDWWNPALRKRA GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLE TGETALHKAACQRNRAVCQLLVDAA TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1			1	1		NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\
CAGPEEQELSLQEEVSTQGTLLESQA GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV  604  1954  A 4548  3 938  QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1	ĺ					EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI
GPQTLQYSYTPQLQDLDPLAQEHTDS EPSTTLVDWDPQTGRLCIPSLSSFDQI PSEGDGLGEEGLLSRLYEEPAPDRPP LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDAA TDSKGKTPQERAQQA\GDPDLAAYTI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1	1	1	[			CAGPEEQELSLQEEVSTQGTLLESQAALAVL
603 1953 A 4543 3 600 YSAVEFVEQASGISDWWPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDAI TDSKGKTPQERAQQA\GDPDLAAYTI YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1			1			GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE
PSEGDGLGEEGLLSRLYEEPAPDRPPI LMQFMEEWGLYVQMEN  603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDAG TDSKGKTPQERAQQA\GDPDLAAYT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSTYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV			1		1		EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE
603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRN GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLD TGETALHKAACQRNRAVCQLLVDAG TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1	1	i				PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY
603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRKRA GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHOPSELLD TGETALHKAACQRNRAVCQLLVDAY TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1		- 1		1		
GMIAPYYEDSDLKDLSHSRVLQSPVS LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHOPSELLD TGETALHKAACQRNRAVCQLLVDAI TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	(00	1050	+-	4543	12	600	VSAVEFVEOASGISDWWNPALRKRMLSDSGI
LQAVIAGDLMKLIESYKNGGSLLIQG LHYAAETGNGEIVKYILDHGPSELLE TGETALHKAACQRNRAVCQLLVDA TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	603	.1933	A	4343	١٠	000	
LHYAAETGNGEIVKYILDHGPSELLE TGETALHKAACQRNRAVCQLLVDA\ TDSKGKTPQERAQQA\GDPDLAA\YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1				1	1	
TGETALHKAACQRNRAVCQLLVDAG TDSKGKTPQERAQQA\GDPDLAA/YT YKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV			1	1		1	I HAY VELCHCEMAKAII DHUDGELI I DWYDGE
TDSKGKTPQERAQQA\GDPDLAA/YTYKVIGHEDLETAV  604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFYSL STAGDPPIPYLTTYGQLSNGDHHFMIQPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV			1	1	}		TCETAL HK A A CODNO A VCOL I VDA GA SI DK
604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1	1	ĺ	ļ	1	1	THEY GET THE A CONTROL AND THE SECOND
604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFEPY QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	]		ı		1	1	
QSNSYPSMSDPYLSSYYPPSIGFPYSL STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	L					<del></del>	
STAGDPPIPYLTTYGQLSNGDHHFMI QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	604	1954	Α	4548	3	938	QUNK VQNGSLHQKD1 VHUNDFEF 1 L1GQAN
QPGGLGNNIYQHRFNFFPENPAFSAV QGQQTQSSAYGSSYTYPPSSLGGTV	1		l			1	
QGQQTQSSAYGSSYTYPPSSLGGTV	1				1		STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG
	1		ì	1		1	QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS
I I I I I I I I I I I I I I I I I I I			- 1	l			QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG
					1		FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA
VKTVGSVVSSVALTGVLSGNGGTNV			1	ſ			VKTVGSVVSSVALTGVLSGNGGTNVNMPVS
			1	}	1		KPTSWAALASKPAKPQPKMKTKSGPVMGGG
LPPPPIKHNMDIGTWDNKGPVPKAP	!	1	ļ	- [			LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP

		- C	Lero	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning		F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	i	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	l	USSN	location	corresponding	1=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	•	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	1	Į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	ł .		peptide		/=possible nucleotide deletion, \=possible
	1			sequence	Ì	nucleotide insertion
	Ļ	<u> </u>	. <del> </del>	Sequence	<del></del>	SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ
	Ĭ	ļ	1		1	OPPQ
	<u> </u>			<u> </u>	0704	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC
605	1955	Α	4553	2	2304	ILLOBKKNCLLMQLEEATKLTSTLQSQLKSLC
	ì	Į.				ASTLTVSSGSSRGSLASSRGSLSSVS
	1		1	1		FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR
	1		1	1	1	DAPPSEPPGPSGFHKQRRSLDTPQSLASLSSRS
	1	1	1	İ		SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE
	1	1	1	1	Ì	GPGLGALDRLRAHASAMGDEDLPGMAALQP
	1	1	l	1	1	HGVPGDGEGPHERGPPPASAPVGGTVTLRED
	1		ł			SAKRLERRARRISACLSDYSLASDSGVFEPLT
	1	1	1	1		KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG
	1	1	1			SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL
	1		Ì			DEL DOCUMENCON AL EDOVE VENEVEDION
	1					PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV
	1	}	1			HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN
		1		1		LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR
			1	1	1	LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR
	j	}	l	)	1	TTAOLOAVERELAEERAKLEYTEEEVLEMER
	Į.	j	1			KEEQAEAISERSWQADSVDSGCSNCTQTSPPY
	1	1		<b>†</b>		PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL
	l	ĺ	1		İ	KVDKETNTEDLFLEEAASLVKERPSRRARGSP
		1				FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS
	1	1	ì	1	1	STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL
	1	1	1			SILPRESPIVENTERRIERIR QUELLOS DE LA PER
	1	1				MARTSLDLELDLQASRTRQRQLNEELCALRE
)	1	Į.	ı	1		LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR
		Ì		1		EAERQTRQTKLDYRHEQAAEKMLKKASKEI
	1	1	1	j		YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL
j	ţ	1	ì		Į.	PADDV
606	1956	A	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP
000	1930	1.7	1333	3	1	VLLLODSSGDYSLAHVREMACSIVDQKFPEC
	1	1	- 1	1	ł	GFYGMYDKILLFRHDPTSENILQLVKAASDIQ
	1	I			1	EGDLIEVVLSASATFEDFQIRPHALFVHSYRA
1	l l	-	1		ļ	PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH
		j		1	}	KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT
	İ			l .	1	SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ
1	- 1	1				SSAELS I SAPPEPLLQASI SESTIONERIO INO CONTROL DE LA CONTRO
1	ľ	1	1	1	1	SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV
1	- 1	- 1	Į.	1		CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA
1		1				PKVPNNCLGEVTINGDLLSPGAESDVVMEEG
1					-	SDDNDSERNSGLMDDMEEAMVQDAEMAMA
	1					ECONDSGEMODPDPDHEDANRTISPSTSNNIP
1		1	1	Į.		LMRVVQSVKHTKRKSSTVMKEGWMVHYTS
1			-			KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE
	1	1		1	1	IPLSEILSLEPVKTSALIPNGANPHCFEITTANV
1		- 1		1	1	I TOPEDODE AKTOURH HOUSELLESS THE
i i	<b>\</b>	-	- 1	1	1	AND THE PROPERTY OF THE PROPER
			ļ	1		VYYVGENVVNPSSPSPNNSVLTSGVGADVAR
						MWEIAIOHALMPVIPKGSSVGTGTNLHRDISV
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCOIQENVDISTVYQIFPDEVLGSGQFGI
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILONLHHPGVVNLECMFETPERVFVVM
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIPPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIPPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDOIONAAFMYPPNPWKEISHEAIDLINN
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLOVKMRKRYSVDKTLSHPWLQDYQTWLDL
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADFFPQV KLCDFGFARIIGEKSFRRSVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADFFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
202	1057		1562		4499	MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL
607	1957	A	4563	1	4499	MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL
607	1957	A	4563	1	4499	MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIPPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQIR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
607	1957	A	4563	1	4499	MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIPPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRCCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
607	1957	A	4563	1	4499	MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIPPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQIR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1				residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
-				peptide	schrence	/=possible nucleotide deletion, \=possible
1		,		sequence		nucleotide insertion
<del></del>		<b> </b>		- sequence		HFNQIETLDPDSFQHLPKLERLFLHNNRITHL
l				}		VPGTFNHLESMKRLRLDSNTLHCDCEILWLA
l	1					DLLKTYAESGNAQAAAICEYPRRIQGRSVATI
	1	1				TPEELNCERPRITSEPQDADVTSGNTVYFTCR
				ł .		AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD
	l'			1		GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ
	ĺ	í		İ	·	EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL
1		ļ				ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS
	1	ļ			ļ	GGLYIQNVVQGDSGEYACSATNNIDSVHATA
	1					FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK
	1	ļ		1		GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI
	ĺ	· ·	1			SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ
]	l .	-				PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP
1		1	1	ł		EPAITWNKDGVQVTESGKFHISPEGFLTINDV
	1		1			GPADAGRYECVARNTIGSASVSMVLSVNVPD   VSRNGDPFVATSIVEAIATVDRAINSTRTHLF
		1	1			DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
1		1			}	ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
1					_	POYLNLIANLSGCTAHRRVNNCSDMCFHQKY
1	1	ţ	(		1	RTHDGTCNNLQHPMWGASLTAFERLLKSVY
	·	1				ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI
		1				GTETVTPDEQFTHMLMQWGQFLDHDLDSTV
		1	1		ĺ	VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN
1		1			1	DSRARSGARCMFFVRSSPVCGSGMTSLLMNS
1	ļ.					VYPREQINQLTSYIDASNVYGSTEHEARSIRD
1	1	1	1	[	1	LASHRGLLRQGIVQRSGKPLLPFATGPPTECM
1		1		i		RDENESPIPCFLAGDHRANEQLGLTSMHTLW
				1		FREHNRIATELLKLNPHWDGDTTYYETRKIVG
	1	1 .	(		1	AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
]		1		1	}	PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD
		1				ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
1					1	RGLFGVAGKMRVPSQLLNTELTERLFSMAHT
		1	1	1	1	VALDLAAINIQRGRDHGIPPYHDYRVYCNLS
		1		1		AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
		1				LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
		1		}		NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
		1			1	LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
			ļ	1		FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS
			}	1	1	A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR
				Ì		TQIKKLESR\LSTTECVDAGGESHANNTKWK
		1		I		KDACTICECKDGQVTCFVEACPPATCAVPVNI
						PGACCPVCLQKRAEEKP
608	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL
			Ì	1		IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL
			1			GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF
				}		YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ
		1	1			LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD
			1	1		EFTIAATDDNHIFAWGNGGNGRLAMTPTERP
1						HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
1	1	1		ĺ		VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE
100	1075	<del> </del>	1.75	<del>                                     </del>	110	GGPDAW
609	1959	A	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
1	1	1	1			PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF
1	1	ļ		}		HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
		1	1	Į.	1	RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
610	1960	1.	4570	697	467	KLSFHVLSGS ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT
010	1900	A	4370	1091	407	GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N
1	1					LVIRPPRPPKVLGLQA
L	ــــــــــــــــــــــــــــــــــــــ	ــــــــــــــــــــــــــــــــــــــ	1			בישיושיושיושי

NO. of marked profits   No. of marked profits   No. of marked   No. of marke						· · · · · · · ·	Amino acid sequence (A=Alanine C=Cysteine,
		SEQ ID	Met	SEQ	Predicted	Predicted end	D-Aspartic Acid E-Glutamic Acid.
Sunce   Sequence   S	NO: of		hod				E=Phenylalanine G=Glycine H=Histidine.
							I=Isoleucine K=I vsine L=Leucine.
1961	eotide		ł				M=Methionine, N=Asparagine, P=Proline,
mainto said residute   residute		uence					O=Glutamine, R=Arginine, S=Serine,
Perilibus of poptible sequence	uence		}	914			T=Threonine, V=Valine, W=Tryptophan,
Poptide			Ì	·			Y=Tyrosine, X=Unknown, *=Stop codon,
1961			l	ľ		acquaise	/=possible nucleotide deletion, \=possible
1961		ì	l	j		}	nucleotide insertion
		1061	<del>  </del>	4571		1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
	611	1961	A	4371	22	1330	WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
VI.ABGILQI.QQG.CANTIGA.HPQSAERAGAN	Ì	1			1		LCAATAVLLSAOGGPVOSKSPRFASWDEMN
LSACGSACQTTGSTDLPLAPESRVDPFVLING   RIQHLQSQFGLDHKHLQDEWARPARENER   ENGALPHONOGILPHKVAQQRRELEQHE   RIQHLQSQFGLDHKHLDDEWARPARENER   ENALPYDPAHNNSKIHRIPRDCQELFQVGER   QSGLFEIDPQGSPFLVNCKMTSDGGWTYUG   RHDGSVDFFRPPWARVAGGGDPTGGFWLGL   REVISITIORINSELAVQLRDWGDNAELLQP   SVILGGEDTAYSLQLATTVPPSG   LSVPFSTWDDMDLRDKNCAKSLSGGWWF   GTCSHSNLNGQYFSSPQORQKLKGIFWKT   WKGRYPFLQATTMLIQPMAAEAAS	ł	1		{		İ	VI.AHGLLOLGOG\CANT\GAHPOSAERAGA\R
LQTQLKAQNSRIQQLFHSYAQQQRRLEKQH  RIQHQSQFGLDHKHLDBEVARPARKRLP  EMAQPVDPAHNVSRIHDLEPDCQLFQUGSP  GSGFEIQPQGSPFUNCKMTSDGGWTVIQR  REDGSVDFNRPWEAYKAGFGDPHGEFWLG  EVAISTGDRNSRLAQLRDWDNAELLQFS  VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG  LSVPSSTWDQDHDLRRDKNCAKSLSGGWWF  GTCSHSNLNGQYFSSPQQRQKLKKGIFWKT  WGRYYPLQATTMLQPMAAEAS  612 1962 A 4575 162 3 FFFETESISSVAGAGVQWRDLSSLQPPPFGSR  GSPASASPVAGTTHHRTRG  613 1963 A 4584 687 321 PLAQRRFFLWVTVKTNGEIBVGSSTYPHFWGS  SNS/RASASVAGIPNARHQARILPVLVEPPF  HTVGRAGLGFLNLAICLPQHFWALGLQAC  HHVGRAGLGFLNLAICLPQHFWALGLQAC  SNS/RASASVAGIPNARHQARILPVLVEPPF  HTVGRAGLGFLNLAICLPQHFWALGLQAC  LNIKPHPAHKYISMIQFNVHFMCMSVHIVI  614 1964 A 4589 727 299 FGSQSAQRGRGRGRRAGASTQTITMSFMG  GGLPCAWVGTILLVVAMATDHWMQYRLSGF  HTVGRAGLGFLNLAICLPQHFWALGLQAC  GGLPCAWVGTILLVVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWMQYRLSGF  GGLPCAWVGTILLVAMATDHWQTTTVSGGP  CWCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	ļ			ì	1	1	LSACGSACOGTEGSTDLPLAPESRVDPEVLHS
RIQHLQSQFGLLDHKHLDHEVARPARKILP   PMAQPVPDAHVSRLHRIPPCQELFQVGER   PMAQPVPDAHVSRLHRIPPCQELFQVGER   QSGLFEIQPQGSPFILVNCKMTSDGGWTVQR   RIDGSVPFIRPWARYKAGPGDPHGEFWLGL   EKVHSITGDRISSRLAVQLRDWDGNAELLQFS   VHLGGEDTAYSWLQTAPWAGQLGATTVPPSG   LSVPFSTWDQDHDLRADKNCAKSLSGGWF   GSSHSNLRGQYRSBPQDQKLKKGIFWKT   WRGRYYPLQATTMLIQPMAAEAAS   VHRGGEDTAYSYRSDGQNGKLKKGIFWKT   WRGRYYPLQATTMLIQPMAAEAAS   SFFETESSEVAQAGQVQWFDSSLQPPPPGSR   GSPASASPYAGGTGTRHRITRG   GSPASASPYAGGTGTRHRITRG   SNSPASASPAGGTGTRHRITRG   SNSPASASPAGGTGRHRATGGACN   LNIKPHPAHKYSMQFNARIFOMSVHTYI   MHYGRAGLGFLJKLALCLPQHFKVLGLQACN   LNIKPHPAHKYSMQFNARIFOMSVHTYI   MHYGRAGLGFLJKLALCLPQHFKVLGLQACN   LNIKPHPAHKYSMQFNARIFOMSVHTYI   GGFCAWVGTILVJVAMATDHVMQYRLSGS   FAHQGLWRYCLGNKCYLQTDSIAYWANTRA   FMISSALCAISGIMGIMAFJWAVJMTFFA   GGFYAWCTULVJVAMATDHVMQYRLSGS   FAHQGLWRYCLGNKCYLQTDSIAYWANTRA   FMISSALCAISGIMGIMAFJWAVJMTFFA   GGFYAWCTULVJVAMATDHVMQYRLSGS   FAHQGLWRYCLGNKCYLQTDSIAYWANTRA   FMISSALCAISGIMGIMAFJWAVJMTFFA   GGFYAWCATRVHECRRLSTFR   THE PRICE OF			1		l	ļ	I OTOLK AONSRIOOLFHKVAOQORHLEKQHL
EMAQPYPPAHNYSRLIRLPROCQELPQVGBR   QSGIFEIQPOGSPFLVNCKMTSDGGWTVIQR   RIDGSVDFNRPWEAYKAGFGDPHGEFWLGL   ERVHSTIGDRISRLAVQLRDWDGNAELLQFS   VYHLGGEDTAYSLQLTAPVAGQLGATTYPSG   LVVPFSTWDDDHDLRADNACASLS.GGWWF   GTCSHSNLNGQYFRSEPQQRQKLKKGIFWKT   WRGRYPJQATTMLIQPMAEAAS   GSPASASPVAGTGTRIHRTRG   FFETESRSVAQAGVQWEDLSSLQFFPFGSR   GSPASASPVAGTGTRIHRTRG   GSPASASPVAGTGTRIHRTRG   GSPASASPVAGTGTRIHRTRG   FHYGRAGLGFLANLACLPCFRF   HHYGRAGLGFRANLACLPCFRF   HHYGRAGLGFRANLACLPCFRF   HHYGRAGLGFRANLACLPCFRF   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLGFRANLACLPCFR   HHYGRAGLACLPCFR   HHY	,		1				RIOHLOSOFGLLDHKHLDHEVAKPARRKRLP
QSGLFEIOPGSSPPFLVNKMTSDGGWYTQR   RIDGSVPPRINPRAYKAGFGDPHGEFWLGL   REVHISTIGRNSRLAVQLRDWDGNAELLQFS   VHLGGEDTAYSQLGATTYPPSG   LEVPSTWDQDHDLRRDKNCAKSLSGGWF   GTCSHSILNGQYFRSIPQQRQKLKKGIFWKT   WRGRYYPLQATIMLIQPMALEAAS   GSPASASPYAGTGTRHHRTRG	İ	į	1				EMAOPVDPAHNVSRLHRLPRDCQELFQVGER
REIDGSVPFNRPWEAYKAGFCOPHOEFWIGL	i	1	1		ì	[	OSGLFEIOPOGSPPFLVNCKMTSDGGWTVIQR
VHLGGDTAYSLQLTAPVAGQLGATYPPSG	1	1	Į.				RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
VHLGGDTAYSLQLTAPVAGQLGATYPPSG	1			1			FKVHSITGDRNSRLAVOLRDWDGNAELLQFS
GTCSHSNLNGQYFRSIPQQRQKLKKGIPWXT	1	1	1	l	1		VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
	I	}	1	ļ		1	LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
1962	l	İ	1		1		GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
GSPASASPVAGITGTRIHIRTEG			1		1		WRGRYYPLQATTMLIQPMAAEAAS
1963	612	1962	1 A	4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
SINSPASASQVAGIPNARIQAGRIIFVILVEPRF	012	1902	1 ^	1373	1		GSPASASPVAGITGTRHHRTRG
SINSPASASQVAGIPNARIQARII VILVEPRE	612	1063	ΙΔ	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
LNIKPHPAHKYISMIGPNYHFMCMSVHIYT	1 913	1303	1.7		1		SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
1964				1	1	1	HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
GGLFCAWVGTLLVVAMATDHWMQYRLSGS FAHQGLWRYCLGNKCVLQTDSIAYWNATRA FMILSALCAISGIIMGIMAF/GWVAVLMTFFA GIFYMCAYRVHECRRLSTPR  615 1965 A 4590 2 414 TILPEKIQAWAQKQCPQSGEBAVALVVHLEK ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHISGHQEQLNRKR ERRPLPKNARPSWYPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGTGYTTVPSOPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAASSFQHSSSLGRELPY DPVDTEGFEGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERRQQNL RASSREHTVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKFLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDVTVSWVWIIGLVIAMA MSLLSIILHLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILISHEVIIILLIFLRKRIL AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRCCPNARCOFAFYGGESGYH RALLGQIFNAFMFWLANFVLALGQVTLAG AFASYYWALKKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCHTCLKCCFWCLEKFIKFLÑRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGLAFFFFTHRIRIVQDT APPLNYYWVPUTLVIVGSVLLAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  ETSVEPPIL GEFRIKLSNNTKVVKTEYKATEY	1		1			``	LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
GGLFCAWVGTLLVVAMTDHWWKLSUS FAHQGWRYCLGNKCVLQTDSIAYWNATRA FMILSALCAISGIIMGIMAF/GWVAVLMTFFA GIFYMCAYRVHECRRLSTPR  615 1965 A 4590 2 414 TILPEKIQAWAQCQPQSGEBAVALVVHLEK ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREPGSLHSGHQGUNKKR ERRPLPKNARPSFWVPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRALANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAASFQHSSSLGRELPY DPVDTEGGFGGDMQERFLFPEYILDPEPQFT REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLANRIFEDVTVSWYWIIGLVIAMA MSLLSIILLHLAGIMGWVMIMEISELGYRE HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLIFLRKRIL AIALIKEARAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGEGGYH RALLGQUFNAFWFWLANFVLALGQVTLAG AFASYYWALKKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCHTCLKCCFWCLEKFKFLÑRNAYIM IAIYGTNFCTSARNAFFLLMRNITRVAVLDKV TDFLFLLGKLLIVGSVGLAFFFFTHRIRNQDT APPLNYYWVPLIVTGSVLLAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  PTSVPEPYIL GEPRIKLSNNTKVVKTEYKATEY	614	1064	- A	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
FMILSALCAISGIIMGIMAFIGWAVLMITFA GIFYMCAYRVHECRRI.STPR TILPEKIQAWAQKQCPQSGEEAVALVVHLEK ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLFKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRVPRLANFVELVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSOPG AGLELPTSGDPPALASQSAGITGVTTVPSOPG SXRHGLREPLLERCAAASSF,GHSSSLGRELPY DPVDTEGFGEGGGMQERFLFFEYILDPEPQPT REKQL,QELQQQQEEEERQRQQREERRQQNL RARSREHPVVGHIPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRFEDYTVSWYWIIGLVIAMA MSLLSILLHLAGIMGWVMIMENSEGLGYBIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSLEVIILLIFLRKBLI AIALKEASRAVGYVMCSLLYPLVTFFLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETTPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFWIFLALGURVLLAGHL AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALLALVQURVLLEVLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLÑRNAYIM IANGTNFCTSARNAFFLLMRNIRVAVLDKV TDFLFLIGKLLIVGSVGILAFFFFTHRRNYQDT APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWVPILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTVIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTYIVGSYLLAFFFFTHRRNYQDL APPLNYYWFILTYLGFFRKLSNNTKVVKTEYKATEY	014	1304	10	1.50	1		GGLFCAWVGTILLVVAMATDHWMQYRLSGS
GIFYMCAYRHECRLSTPR  GIFYMCAYRHECRLSTPR  TILPEKIQAWQKQCPQSGEEAVALVVHLEK ETGRLRQVSSPVHREKHSPLGAAMEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFYFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG AGLELPTSGDPPALASQSAGITGVTTVPSGPG AGLELPTSGDPPALASQSAGITGVTTVPSGPG DPVDTEGFGEGGDMQERFLFPEYLDPEPQPT REKQLQELQQEEEERQQQREEERRQQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIMAA MSILSIILHLLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHRQTWLAFMILSILEVIIILLIFLRKBLI ATALKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALIGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLÑRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVOSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFFTHRRIVQDT APPLNYYWYPILTVIVGSVGILAFFSTYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES PTSVEPPYILGFFRYKLSNNTKVVKTEYKATEY	1			l	1		FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
615 1965 A 4590 2 414 TILPEKIQAWAQKQCPQSGEBAVALVVHLEK ETGRLRQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREPGSLHSGHQEQLNRKR ERPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFVFLVEMGFLHVGQ AGLELPTSODPALASQSAGITGVTTVPSGPG AGLELPTSODPALASQSAGITGVTTVPSGPG REVOLOGITATIVPSGPG REV			-		1		FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
616 1966 A 4592 773 488 DFALVAGAGVVMNLGSPQPLPPGFKRFSCL  616 1966 A 4592 773 488 DFALVAQAGVWDHNLGSPQPLPPGFKRFSCL  SLPSSWYRCVPPRLANFVFLVEMGFLHVGQ  AGLELPTSGDPPALASQSAGITGVTTVPSGPG  617 1967 B 4595 84 478 XRRGLREPLERRCAASSFQHSSSLGRELPY  DPVDTEGFGEGGDMQERFLFPEYTLDPEPQPT  REKQLQELQQQEEEERQRQQRREERRQQNL  RARSREHPVVGHPDPALPPSGVNCSGCGAEL  HCQDAR*  ARSRNSARGVYGMCVDTLFLCFLEDLERNDG  SAERPYFMCSTLKKPLARRCFPAIHAYKGVL  MVGNETTYEDGHGSRKNITDLVEGAKKANG  VLEARQLAMRIFEDYTVSWYVMIIGLVIAMA  MSLLSIILHILAGIMGWVMIIMEISELGYRIF  HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV  YLHLRQTWLAFMILSLEVIILLLIFLRKRILI  AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI  AYWASTAVFLSTSNEAVYKIFDDSPCFTAKT  CNPETFPSSNESRQCFNARCOFAFYGGESGYH  RALIGLQIFNAFMFFWLANFVLALGQVTLAG  AFASYYWALRKPDDLPAFPLFSAFGRALRYH  TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN  KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM  LAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV  TDFLFLLGKLLIVGSVGILAFFFFTHRIRVQDT  APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC  VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL  LNKTNKKAAES			Ì		1		GIFYMCAYRVHECRRLSTPR
ETGRLRQQVSSPVHREKHSPLGAAW ADPQ PEQVETOPRAVSREEPGSLHISGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWYRCVPPRLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG SLPSSWYRCVPPRLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG PPUTDEGFGEGGDMQERFLFPEYTLDPEPQPT REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHPVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWHIGLVIAMA MSLLSIILHILAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSLEVIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFELCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPTAKT CNPETFPSSNESRQCFNARCOFAFYGGESGYH RALIGLQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	615	1965	A	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD  616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG  617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHVYGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGMETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILHILAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIILLLIFLRKILI AIALIKEASRAVGYYMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETTPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRYDDLPAFFLFSAFGRALRYM TGSLAFGALILAIVQIIRVILEYJDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LINKTNKKAAES ETSVEPFVI.GEPRKLSNNTKVVKKTEYKATEY	013	1505	1	1.55	_		ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
TRLPAGSQEPVKD  FALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG  AGLELPTSGDPPALASQSAGITGVTTVPSGPG  XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  ASSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILHLLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHRQTWLAFMIILSILEVIIILLLFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM LAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRRIVQDT APPLNYYWPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  ETSVFPYILGFF/RKLSNNTKVVKKTEYKATEY	-	ļ				Ĭ	PEQVETQPRAVSREEPGSLHSGHQEQLNRKK
616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG AGLELPTSGDPPALASQSAGITGVTTVPSGPG XRHGIREPLLERCAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQEEEERQRQQREERRQQNL RARSREIPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRGFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHRQYMLAFMILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTIFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNTCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVGDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVGDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVGDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVGDT APPLNYYWVPILTVIVGSVGIAFFFFTHRIRIVGDT APPLNYYWFTHATAGT ARTTER ARTTER APPLNYTH ATTER ARTT		1	ł			1	ERRPLPKNARPSPWVPALADEWNILHQEVII
SLPSSWEYRCVPP/RLANFVFLVEMGELHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG AGLELPTSGDPPALASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQREERRQQNL RARSREHPVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHIAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILHHLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTIFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMIFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLÑRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  PTSVEPYULGFFRKLSNNTKVVKTEYKATEY			1	İ			TRLPAGSQEPVKD
SLPSSWEYRCVPYRAMINEVENDENDENDENDENDENDENDENDENDENDENDENDENDE	616	1966	A	4592	773	488	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGMMQERFLFPEYILDPEPQPT REKQLQELQQQEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMERISELGYRIF HCYMGYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETTPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	010	1,700			-		SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILHHLAGIMGWVMIIMERISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLÑRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES RTSVFPVILGEF/RKLSNNTKVVKTEYKATEY							AGLELPTSGDPPALASQSAGTTGVTTVPSGPG
DPVDTEGFGEGGDMQERFLFFEYILDFEFQFT REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHILLAGIMGWVMIIMERSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	617	1967	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELFY
RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEINSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  PTSVEPYILGEF/RKLSNNTKVVKTEYKATEY	0.7	170.	-				DPVDTEGFGEGGDMQERFLFPEYILDPEPQF1
HCQDAR*  618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWPILTVIVGSYGILAFFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES		1	-				REKQLQELQQQQEEEERQRQQRREERRQQNL
618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFFAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCFFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			-				
SAERPYFMCSTLKKPLARRCFFAIHAYKGYL MYGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWPILTVIVGSYGILAFFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							HCQDAR*
SAERPYFMCSTLKRPLARRCFFAIHAY KUYL  MYGNETTYEDGHGSRKNITDLVEGAKKANG  VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA  MSLLSIILHILAGIMGWVMIIMENSELGYRIF  HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV  YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI  AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI  AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT  CNPETFPSSNESRQCPNARCQFAFYGGESGYH  RALLGLQIFNAFMFFWLANFVLALGQVTLAG  AFASYYWALRKPDDLPAFPLFSAFGRALRYH  TGSLAFGALILAIVQIRVILEYLDQRLKAAEN  KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM  IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV  TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT  APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC  VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL  LNKTNKKAAES	618	1968	A	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILHILAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHI.RQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYYMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCOFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFILMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIIVQDT APPLNYYWVPILTVIVGSYGILAFFFFTNIRIVQDT VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	"."		"-			i	SAERPYFMCSTLKKPLARRCFFAIHAYKUVL
MSLLSIILHLLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCOFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIIVQDT APPLNYYWVPILTVIVGSYGILAFFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			Į		· <b> </b>	l	MVGNETTYEDGHGSRKNITDLVEGAKKANG
HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	1		1	ì	1	1	VLEARQLAMKIFEDYTVSWYWIIIGLVIAMA
YLHLRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTIFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES		j		i		1	MSLLSIILLHLLAGIMGW VMIIMEISELGYKIF
AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCFFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALLAIVQIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			- 1	1	1	1	HCYMEYSKLKGEAGSDVSLVDLGFQ1DFKV
AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			1	]	- 1	1	YLHLROTWLAFMILLSILEVIIILLLIFLRARILI
CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVFPYILGEF/RKLSNNTKVVKTEYKATEY	- [		ł	- 1			AIALIKEASKAVGYVMUSLLYPLVIFFLLULUI
RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY	i				(		AYWASTAVILSTSNEAVIATIOUSICITIANT
AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY			1	i	1	1	CNPETFPSSNESKQCPNAKCQFAF I GOETH AC
TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY	1		1	1			KALLGLQIFNAFMIT WLANT YLALGQV ILAG
KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY	-	1	1	1	1		AFASYYWALKKYUULPAPPLISAFUKALKIII
IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY	1			1			TGSLAFGALILATVQHKVILEYLDQKLKAAEN
TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES BTSVEPYILGEF/RKLSNNTKVVKTEYKATEY	1		l l	1			KFAKCLMTCLKCCF WCLEKFIKFLNKNA I IVI
APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES  RTSVEPYILGEP/RKLSNNTKVVKTEYKATEY	1		ļ	1	l l	1	IAIYGINFCISARNAFFLLMRNIRVAVLDKV
VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES RTSVEPYILGEP/RKLSNNTKVVKTEYKATEY		1	- 1	1	1		TDFLFLLGKLLIVGSVGILAFFFFTHKIKIVQD1
LNKTNKKAAES  257 RTSVEPYILGEP/RKLSNNTKVVKTEYKATEY			1	- 1	1		APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
RTSVEPYÜ GEF/RKLSNNTKVVKTEYKATEY			1	1.		. ]	VDTLFLCFLEDLERNDGSAERP 1 FM551LKKL
619 1969 A 4601 2 357 RTSVEPYILGEF/RKLSNNTKVVKTEFRATET			i				LNKINKKAAES
	619	1969	A	4601	2	357	RISVEPYILGEF/KKLSNNIKVVKIEIKAIEI

					The affect of and	Amina anid seguence (AmAlanina CoCustaina
	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
	NO: of	hod	ID NO:	beginning	nucleotide	D. Physiologica C. Chaine H. Wistidine
	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
- 1		-		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
. 1			,	peptide	•	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
+				30queneo		GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK
						GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN
		1		<b> </b>		NOHVECNEICHRLSLTRPSMEKPCKS
				<u> </u>		
620	1970	Α	4606	1	2415	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR
		l	l	ì	1	KGHLEEEEEDGEEGAETLAHFCPMELRGPEP
j			)		Į.	LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF
			}		ļ	TGAFLLGYVAFRGSCQACGDSVLVVSEDVN
i f		1	ļ		ŀ	YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
i l		ĺ	Í	{	[	EDTIRQTSLRERVAGSAGMAALTQDIRAALS
( <b>!</b>		Ì	Į.	ļ		RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV
1 1		1		1	1	DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
			1	{	Į	VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
( 1		l		ļ	J.	GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ
		,		1		DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
i i		l	l			
1 1		1	1	1	1	NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
1		l		ł		KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
1 1		ŀ	i	1	1,	NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA
i i		[	1	ł		WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
<b>!</b> !		į,		i	}	-PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
[ [		{		1	ļ -	HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
1 1			1		<b>,</b>	IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
1	}	1	1	1	1	AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\
		1	1			DDO\AYPFLHTKEDTYENLHKVLQGRLPAVA
		J		1		QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV
ļ		1	Į.			LRHIGNLNEFSGDLKARGLTLQWVYSARGDY
	l					IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV
1 1		1	1			
	1	1	1	1		EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL
] !		1	1		ł	DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\
\		<u> </u>				ALL\TWDACKGAANALSGDVWNIDNNF
621	1971	Α	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\
1	1	1				NTLVLKQQTFIESARSIGASDMTVLLRHILPGT
1	ł	1	ļ			GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP
1	ł	1	1	l .		EWGAMLNEARADMVIAPHVAVFPALAIFLTV
j '			İ			LAFNLLGDGLRDALDPKIKG
622	1972	A	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
022	13/2	^	1017	~	1	CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF
	!	1	<b>!</b>	1	}	RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
1	1	(		Į.		SCVILLGLLLYDVFFVFITFITKNGESIMVEL
		1	1	l _		
1	}	1	1		1	AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV
	!	1	1			VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL
	1	1	1	}		LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL
1		1	1	1	· ·	TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV
		1	1	1	1	AWETVREMKKFWERVTS
623	1973	A	4619	17	691	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF
	1	1		1		GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT
1		1	1	1	1	DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ
1			1		1	IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
1		1	1		1	OYFDSRGMFISIVFSAPLLVNAMIIVVMWVW
	1	1	ļ			KTLNVMTDLKNAQERRKEKKRRKED*GAA
}	1	1	1	Į	1	
	1	1	1			AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG
						LKNRCFI
624	1974	A	4622	164	668	VSCYTALQSIMNQPESANDPEPLCAVCGQAH
	1		1	ĺ		SLEENHFYSYPEEVDDDLICHICLQALLDPLD
	1	1	1	ł		TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL
1	1			1	1	OHCKKSSILVNKLLNKLLVTCPFREHCTQVL
1	1		1	}		ORCDLEHHFQTSQAWGTHL*SQLLGRLRQED
	ı	ļ	1	1		CLSPGVHHCSEV
	i	1				
(22	13055	<del>  </del>	1400	174	172	
625	1975	A	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP PPLLIPSS*LSP

CEO 110	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Asnartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	sed-	{	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Lacitor	ĺ	914	ng to first	acid residue	O=Ghrtamine, R=Arginine, S=Serine,
ucnee.	ļ	}	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI
		ļ			}	LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP
	1	-	·	i		ASASQVAGIAGTHH
627	1977	A	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK
	İ	<u> </u>				QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP
	İ	Ì				PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP
						ARAYL
628	1978	Α	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK
			1			NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF
					ľ	YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK
	1	ĺ	]		Į.	VFYICFTEFLLFLYFL*LFIIKVSCSH*CSTICVF
		1 .	1		ļ	SYKSFAVIIFFVDNTRFFSFGF
		<del> </del>	1660	10	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV
629	1979	A	4660	18	797	KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH
		1		İ	1	PKLVFSQEGRYVKNTASASSWPVFSSAWNYF
	1		į	İ		AGWRNPQKTAFVERFQHLSCVLGKNVFTSG
		1				KHYWEVESRDSLEVAVGVCREDVMGITDRS
		1		i		KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ
		1				EPALHRVGVYLDRGTGNVSFYSAVDGVHLH
			ļ			TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G
			j			FSSPDONSFPVVQLRDTHPWALFCPSCLYPG
		1	1			WSIFWVSLTVPFGICPLCASQEAVPWEVGLA
	1	1				NGDGTGNFPRRFWEIFL
630	1980	A	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF
050	1700			1	1	TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG
1				1		FHDVGQDGLDLLTS*STPSASQSAEITGVSHC
1	1	1			l	TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV
		1	ļ	i		AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ
1	-		1			NPVFLERRPRALHSSPGLTTQRILWAQGLWV
		-	1	1		GAGSTGCSRGPRGEGVFREG
632	1982	A	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP
i		1	1	Í		*LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP
						ASGLAPVPTHWTVSELSRSPVATATFC
633 ·	1983	Α	4696	1	1365	RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS
		Ì				KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP
1		1		l .	1	DLKDLFITVDEPESHVTTIETFITYRITKTSRG
						EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII
Ì			1		•	PPLPEKFIVKGMVERFNDDFIETRRKALHKFL
1	İ				Į	NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ
1			1		1	GPGLLSRMGQTVRAVASSMRGVKNRPEEFM
1		ļ	ı			
			1			
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT
	1004		4700	421	158	EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SOTNLHLEEASEDKP
634	1984	A	4708	421	158	EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ
634	1984	À	4708	421	158	EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH
634	1984	A	4708	421	158	EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH YIKQPDAKERRRTVHWKKETESEASEITIPPST
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH YIKQPDAKERRRTVHWKKTESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT

WO 01/57188

			T	<del></del>	D Card and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence	ĺ	[	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	Į	peptide	Soques	/-possible nucleotide deletion, \-possible
	l	l		sequence		nucleotide insertion
	<u> </u>		<del> </del>	sequence	<del> </del>	ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
	1	1				MYFTTPSNHNAYOVDSVQST
(20	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
637	1987	A .	7/20	1007	1 233	LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS
	1		1	Į.	Į	DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
		1	Ì	1		HAG*AGLELLTSGDPPASASRSAGITGVSHHA
		1		į.		RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
030	1700	1 **	17737	1	1	TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
	1	į			1	YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL
	1		ĺ	Ì	i	GOYTSOGGVTAWRKICPIFEGIGYASQMIVIL
	ļ	1	-			LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
		1	ļ			WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	TA-	4743	1040	699	OGLTLLPRMECSATITAHCSLELPGSIDLPTSA
039	1707	11	1	1.0.10	1	S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
	1	1		<b>!</b>	Į.	AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA
,			1	1		TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
040	1,550	1			ł	WIRRRPCLPSGCLKMNREIGPLQHSLCCPGWS
ļ	]					QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
	1	1		1	1	MARSRLTATSASQVQAILLPQPPGTTDSCSPS
	Į.	1	l	ì	1	PDHEQQPLSWVLPPPQKDMNPREQQVALGP
ļ						QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
1		1			1	LQLAASPYFSPSWAECPQPVPAGTHATWCLA
1	1	[	ı	[		RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
		1				FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
1						QRDTDTGVHTGSGTHTHAHTPPEK
642	1992	A	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
		}	1	1	<b>{</b>	FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
	1		j			TWWFGVKFAAGGLGTFHALLNTAVHVVMY
		1				SYYGLSALGPAYQKYLWWKKYLTSLQLVQF
		1	}	j	1	VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
	ļ	ł	1			FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
		j		-		QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
i	1	}		ļ		MVYFVGENNGDSSHNPVLAATGVGLDVAQS
1		Ì		1	1	WEKAIRQALMPVTPQASVCTSPGQGKDHSK-
1		1			_	Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
1						LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
		-	1	1		AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
L	_L				1	SYKDIWGWPCLCGVLHAYIPLLV LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV
645	1995	A	4805	458	126	AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
1		]		1	1	PLLAGLVAADAVASLLIVGAVFLCARPRRSP
1		l			i	
L					1600	AQEDGKVYINMPGRG LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
646	1996	Α	4817	47	1033	LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
1	Į.	1				HKQAVQCLKGPGQVARLVLERRVPRSTQQC
1	1	1	1			PSANDSMGDERTAVSLVTALPGRPSSCVSVT
1		1	1		1	LOUND HOLD THE LOUIS HER L
		-		1	1	DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL
i	1	1		1	1	KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE
	1	İ				WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
1	1	1			1	YYPAAVEVLHLLRGAPQEVTLLLCRPPPGAL
ļ	l l	ì	<b>\</b> .	- 1		PELEQEWQTPELSADKEFTRATCTDSCTSPIL GSRGQLGGTVPPQMQGKAWGLRPESSQKAII
	l	İ	\ \ \ \	1		EGTMGAKTERDLGPVP
					<del></del>	PRVRGDWPLEKKKSNSNIHPIPSWCGSTDSKI
647	1997	A	4854	1044	335	LECANDIDALPEVIVORISMILLING A COSTOCAT

Seq   D   No. of   nucleoted				Long	15 11 1	Valuation and	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide   sequence   SSN   09496   914   1   1   1   1   1   1   1   1   1	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A-Alainne C-Cysteine,
Section			hod				E-Phenylalenine G-Glycine HaHistidine
			ŀ				I-Isolausine K-I veine I al encine
1914   1915				1			
	seq-	uence	l				
Peptide   Sequence   Peptide   Sequence   Peptide   Pe	nence	ļ	}	914			Q=Glutamine, K=Arginine, S=Serille,
peptide		į		l .			
		1	1	ļ		sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
		1	}			]	
GPPWESKNSTAVWRGRDSKERLELVKLSRE   HPELIDAAFINFFFEKBEDNLOGPIVKHISPED   FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK   QDSTYSHEFYNELQWKHYPYKSKI SDLLEK   LKWAKDHDEEAKKIAKAGQFARNNIMGD   DDCYYFQTFPRMPYKY   AGMLPAVGSADEEDFAEDCPELVPMETTQ   GDELYCTTIPTMITGYLGAGKTTLINYI   LTEQHSKRVAVILNEFGGGALEKSLAVSQG   GELYEEWLELRNGCLCSVKDNGLRAIENIM   QKKGKFYILLETTGLADPGAVASMFWDA   ELGSDIYLDGHITVDSKYGLKHIAEBEPDOLI   NEATRQVALADALLNKTLUVFEEDKALEKSLAVSQG   GELYEEWLELRNGCLCSVKDNGLRAIENIM   QKKGKFYILLETTGLADPGAVASMFWDA   ELGSDIYLDGHITVDSKYGLKHIAEBEPDOLI   NEATRQVALADALLNKTLUVFEEDKYALEKSLAVSQG   GELYEEWLELRNGCLCSVKDNGLRAIENIM   QKKGKFYILLETTGLADPGAVASMFWDA   ELGSDIYLDGHITVDSKYGLKHIAEBEPDOLI   NEATRQVALADALLNKTLUVFEEDKYALENIM   QKKGKFYILLETTGLADPGAVASMFWDA   ELGSDIYLDGHITVDSKYGLKHIAEBEPDOLI   NEATRQVALADALLNKTLUVFEEDKYALENIM   NEATRQVALADALLHAEDDLKQLFIAT   TIRSINGLQQLETQRSVDLSDFLACHSTLUTTUPOQ   NAKEEHLINMTIQNLLWEKNYRKDNHCAEV   VYETEKQWTHRKEDQVCT   VYETEKQWTHKEDQVCT   VYETEKQWTHKEDQVCT   VYETEKQWTHKEDQVCT   VYETEKQWTHKEDQVCT   VYETEKGWTHKEDQVCT   VYETAKGT   VYETEKGWTHKEDQVCT   VYETAKGT   VYETEKGWTHKEDQVCT   VYETAKGT   VYETEKGWTHKEDQVCT   VYETAKG					sequence		
HPELIDAAFTNFFFKHDENLYGPIVKHISFFD   FKKKYQNINGTVAAYRLYPLYGSVVLK   QDSIYYEHFYNELQPWKHYPYKSNLSDILES   LWWAKNHDEEAKKIAKAGGEFARNILMGD   DIFCYYFGTFRNNPTYK   AGMLPAVGSADEEDPAEEDPAEEDPAEEDPAEEDPAEEDPAEEDPAEE							IVMPTYDLTDSVLETMGRVSLDMMSVQANT
FFKIKKYQNINGTYAAYRLPYLLVGDSYVLK	)	1		1		1	
GSS				1			HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD
GSS			į.	1	İ		FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK
LKWARDHDEAKKIAKAGGEARNNLMGD   DIFCYPRGTPRINPYK	ļ	1	i .		l		
	1		1	· l			LKWAKDHDEEAKKIAKAGQEFARNNLMGD
648	í	1	ſ				
SEEEKSGLGAKIPYTIIITGYLGAGKITILLNY    LTEQHISKRVAVINERGEGGSLSLAVSOG    GELYPEWLELRNGCLCCSVKDNGLRALENIM   QKKGKEPDYLLETTGLADPGAVASMFWVDA    ELGSDIYLDGIITIVDSKYGLKHLAEBKPDGLI    NRATRQVALADALINKYDLVPEDVKKLRI    TIRSINGLGQILETQRSRVDLSNYLDLHAFDSL    SGISLQKKLQHVPGTQPHLDQSIVTITDYPG    NAKEREHLNMFQNLLWEKNYNRKDNHCMEV    SUKDTERTINFIVLLIGRINLDKIKQLFIAT     TYTETEKQWTHFKEDQVCT     SWEDTERTINFIVLLIGRINLDKIKQLFIAT     TYTETEKQWTHFKEDQVCT     OGVSLLIPKLGYWAQYWAHWQPPLPGFKR     FSCLSLRSSWP*KCAPPHPAFVFLVEMGFHRV     QAGLELRTSGDPASASQSAGITGVFHRVE     GAGLERTSGDPASASQSAGITGVFHRVE     GAGLERTSGDPASASQSAGITGVFHRVE     GAGLERTSGDPASASQSAGITGVFHPFF     FFLRRSFAFVAQAGVWCDLGSPQPLPFGF     K*PSCLSLRSSWDYRHAPPPCP*FLVF**RQQ     FTMLARLVINS*PHDLFTSPSQSAEIKGVSHR     CPASFYLFLKYYLEAKFCA*GPSAVCAA     GYRGHKSCLLINCVVQI     DAWGPETRLARLINFDSFIEPRFGRLPELEATR     PHMEPKASCPAAPLINFESFIEPRFGRLPELEATR     PHMEPKASCPAAPLINFESFIEPRFGRLPELEATR     PHMEPKASCPAAPLINFOSFIEPRFGRLPELEATR     PHMEPKASCPAAPLIFCAMNTAMWEHPITAQQ     VTQLKAFGYVEIPCVAKKLVCGDEGLGAMA     EVOTITORVKWELFGHSGFQQS*PGISVMGVP     LYSEWVQAKSVKMDVGKIGGYPHLINGGPA     LSLPRQACSRLNWTEGFGLGQS*PGISVMGVP     LYSEWVQAKSVKMDVGKIGGYPHLINGGPA     LSLPRQACSRLNWTEGFGLGQS*PGISVMGVP     LYSEWVQAKSVKMDVGKIGGYPHLINGGPA     LSLPRQACSRLNWTEGFGLGQS*PGSSCCPA     SP*LPCASNRLAFGGLIFFCAMTAMWEHPITAQQ     VTQLKAFGYVEIPCVAKKLVCGDEGLGAMA     EVOTITORVKWELFGHSGFQQS*PGISVMGVP     LYSEWVQAKSVKMDVGKIGGYPHLINGGPA     LSLPRQACSRLNWTEGFGLGGAMA     GVANCERLPSFWQFTAMAWSWLTATASRR     CPASSRLSSPCSWAQPGVALARCAGVC     KPGDSWRVAACISGRCCSRRGSGSPRNPE     GSTGAWQPSFWGSWKSQRELSAGGAQAW     LLGAGSGLGGA     LSLPRGACSRLNWTEGFGLESLATTEGO     BPAGLGALGVIPVRVPQRPTQRSQGRGW     DPERDPCCRVQVSRGRRGGRKTFGLGCL     PPCLTHLAAACSVCVVWCGRWKRDSAECCCD     BPAGLGALGVIPVRVPQRPTQRSQGRGW     DPERDPCCRVQVSRGRRGGRGKTFGLGCL     PPCLTHLAAACSVCVVWCGRWKRDSAECCCD     BCSAVSQQEBRCRSSSCS     SPFLTHLARLLVLSHLVLSHLVLSHLVLSLU     SPVHKTMR*HELPRLEKKNNNKD     SPVHKTMR*HELPRLEKKNNNKD     SPVHKTMR*HELPRLEKKNNNKD     SPVHKTMR*HELPRLEKKNNNKD     SPVHKTMR*HELPRLEKKNNNKD	740	1000	<u> </u>	1067	2020	937	AGMI.PAVGSADEEEDPAEEDCPELVPMETTO
LITEQHISKRYAVILNEEGGSALEKSILAVSQG   GELYERWLERNGCLCSYAVDA   GELYERWLERNGCLCSYAWDA   ELGSDIYLDGIITIVDSKYGLKHAERKPOGLI   NEATRQVALADALINKTDLVPEEDVKKLRI   TIRSINGLGQILETQRSRYDLSNYLDLHAEDSL   SGISLQKKLQHVROTQPHLDGIVTITDVPG   NAKEHLINMFQINLL WEKNYRNKDNHCMEV   IRLKGLVSKIKSKSQQVIVQGVHELYDLEETPV   SWKDDTBRTNRLVLLGRILDKDLKQLFIAT   VITERKQWTIHFKEDQVCT	648	1998	I A	460/	2030	037	SEEEEKSCI CAKIPVTIITCVI CACKTTI I NYI
GELYPEWLERNGCLCSYNDNGLRAENM   QKKGKFDYLLETTGLADFGAVASMFWVDA   ELGSDIYLDGITTVDSKYGLKHLAEREKPDGLI   NEATRQVALADALINKTDLYPEEDVKKLRT   TIRSINGLGQILETQRSRYDLSNYLDLIAFDSL   SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG   NAKEEHLNMFIQNILWERNYRNKDNHCMEV   RILKGLVSKDKSQQVIVQGVHELYDLEETFV   SWKDDTERTIRVLLGRILDKDILKQLFIAT   VTETEKQWTTHFKEDQVCT   GVGALELETSGDPFASASQSAGTGVSHLA*P   TVETTEKQWTTHFKEDQVCT   GVGALELETSGDPFASASQSAGTGVSHLA*P   TSMPLLPFQRLCYYI   TSMPLLPFQRLCYYI   TFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF   FSCLSLRSSWD*KCAPPHPAFVFLVEMGPHRV   GQAGLELRTSGDPFASASQSAGTGVSHLA*P   TSMPLLPFQRLCYYI   TFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF   FFSCLSLRSSWDYRIAPPPCPS*FLYF**RQQ   FTMLARILVINS*PHDLTTSFSQSAEIKGVSHR   CPASFYLFLKYYLEAKFCA*GPSAGVAG   GYKRGHKSCLLINCVVQI   GAMGPETILARILNYDSFIEPRFGRILPELATR   PHMEPKASCPAAPLMEKFHVLVGVTGSV   AALKLPLJVSKLDDGLGLAVTTERAKHFY   SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL   RRWADLLLVAFLDASHTLAKVASGICDNLTC   VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVGGDEGLGAMA   EVGTIVDKVKEVLLCPGBGROYCHTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVGGDEGLGAMA   EVGTIVDKVKEVLLCPGBGROYCHTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVGGDEGLGAMA   EVGTIVDKVKEVLLCPGBGROYCHTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVGGDEGLGAMA   EVGTIVDKVKEVLLCPGBGROGSPROFTSGEAA   SPPLPCASNRLAFGGLIPPCAKAGVC   KPGDSWLVAKACISGCCSRGRRGGSGPRNFSGECPA   SPPLPCASNRLAFGGLIPPCAKAGVC   KPGDSWLVAKACISGCCSRGRRGGSGPRNFSGEPA   PGFTGRGGFTWAGVSPROFTWGPSFVGFTAGST   VGAVSCRLIPSSWDVRHATMG*PF*YF**R   WGFTLALVLNS*PQVICPPWPKVLTLQA   PGFTGRGGGFT   VGAVSCRLIPSSWDVRHATMG*PF*YF*YF**R   WGFTLALVLNS*PQVICPPWPKVLTLQA   PPCLTHLAAAGCVVVVCQGKKRDSAGQCD   PPCLTHLAAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PPCLTHLAAGCVVVVCQGKKRDSAGCQCL   PCLTHLAGGCGCC   PSCLTHAGCCCCC   PSCLTHAGCCCCCCCCCCCCCCCCCCCCCC		1	ł	1		1	TECHENDAMI VIII OTECHERAL AVSOR
OKKGKFDYILLETIGLAPGAVASMFWVDA	1	ļ	1		1	1	CIE ABEMI EI BRICCI CCEANDROI D'AIEM W
ELGSDIYLDGIITIVDSKYGLKHLÆKEKPOGLI   MEATRQVALADALINKTÜLVPEEDVKKLRT     TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL     SGISLQKKLQHPYGTQPHLDQSIVTITIFDYPG     NAKEEHLNMFIQNLL WEKNYRNKDINICMEV     BLKGLVSIKDKSQQVIVQCYHELYDLEETEV     SWKDDTERTKNEVLLGRINLDKDILKQLKJFIAT     VTETEKQWTHFKEDQVCT     OF VILLERLOV     OF VILLERLOV     SVKDDTERTKNEVLLGRINLDKDILKQLKJFIAT     VTETEKQWTHFKEDQVCT     OF VILLERLOV	l	1	1	1	1	ļ	
NEATRQVALADALINTDI/PEEDVKLIRT   TIRSINGLGQQLETORSRYDLSNVLDLHAFDSL   SGISLQKKLQHVPGTOPHLDQSIVTITTDVPG   NAKEHLNMFIQNIL WEKNVRNKDNHCMEV   RIKGLVSIKDK SQQVIVQVHELYDLEETIV   SWKDDTERTNRLV1LGRNLDKDLKQFIAT   VTETEKQWTHFKEDQVCT	1	1 ,	l	1	1	1	
TRISINGLGQUETQRSRVDLSNVLDIHAPDSL   SGISLOKKLQHVPGTOPPLDGSIVTTIFDVPG    NAKEHLNMFIQNLLWEKNYRNKDNHCMEV    RLKGLVSIKDKSQQVIVQQVHQLETPV    SWKDDTERTNRLVLLGRNLDKDILKQLFIAT     VTETEKQWTTHFKEDQVCT     OF VILLER   VTETEKQWTTHFKEDQVCT     OF VILLER   VTETEKQWTTHFKEDQVCT     OF VILLER   VTETEKQWTHFKEDQVCT     OF VILLER   VTETEKQWTHFKEDQVCT     OR VILLER   VTETEKQWTHFKEDQ						ŀ	ELGSDIYLDGIIIIVDSKYGLKHLAEEKPDGLI
SGISLQKKLQHVPGTOPHLDQSIVTITFDVPG   NAKEHLAMPIONIL KENVYNKKDNECMEV   RLKGLVSIKDKSQQVIVQGVHELYDLEETPV   SWKDDTERTNRLVLLGRNLDKDLKQLFIAT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   VTETEKQWTIHFKEDQVCT   SVSLLLPKLGVQWAQYWAHWQPFLOPFKER   FSCLSLRSSWD*KCAPHPAFVFLVEMGFHRV   GQAGLELRTSGDPPASASQSAGITGVSHLA*P   TSMPLLPFQRLCVYI   FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF   K*FSCLSLPSSWDYRHAPPCPS*FLYF**RQQ   FTMLARLV.NS*PIDLPTSPSQSAEIKGVSIR   CPASFYLFLKYYLEAKFCA*GECAPSAGVGA   GYKRGHKSCLLINCVVQI   DAWGPETRLARLINDSFIEPRPGRLPELEATR   PHMEPKASCPAAAPLMERKFHVLVGYTGSV   AALKLPLLVSKLIDPGLEVAVVTTERAKHFY   SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL   RRWADLLLVARLDANTLGKVASGICDNLLTC   VMRAWDRSKPLLFCAPMNTAMWEHPTTAQQ   VDQLKAFGYVEPCVAKKLVCOEDGLGAMA   EVGTIVOKVKEVLFOHSGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKUGGYPHLLNGGPA   LSLPRGQACSRLNWTEGPCLSFFQGEAAA   SP*LPCASNRLAFGGLIFPCAPLVYPAPFSPLL   AFSCAPPRAHTISRTHPSAPLVVPKPSSRAR   GQSPJPSRASSPSSWAQVPGVALARCAGVC   KPGDSWRVAACISGRCCSGGRRGSGPRNPE   QSFRCAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGGA   LLGSAGSGLRGGA   FFFFFFGVSVLCHEGWTAVARSWLTATSASR   VQAYSCFRLPSSWDYRHATMPG4FF*YF**R   WGFTLALLVLNS*PQVICPPWPPKVLTLQA   RFGTFGREFRRSWFQCLPPWFPFGRSGGRB   PAVGLGALGVUPPVRYPQRSQCRGW   DPENDFGCRVOVSKGRFRGGGKPTGQSCLGCL   PPCLTHLAAASCVVVWGGRWKDSAECQCD   HSCSAVSQEDRCRSSCC   HSCSAVSQEDRCRSSCC   HSCSAVSQEDRCRSSCC   CVFICTION   TLSQWIPLTKIJGKKISTHYLSHLVTANLLVC   CVFICTION   TLSQWIPLTKIJGKKISTHYLSHLVTANLLVC   CVFICTION   TREWHIKTMR*HIPLPRLEKKNNIKD	1	1	1	i		1	NEATRQVALADAILINKTDLVPEEDVKKLKI
NAKEEHLNMFIQNLLWEKNYKKDNHCMEV   RILKGLVSKIKDKSQUVGVERLYDLEETPV SWKDDTERTNRLVLLGRNLDKDILKQLFIAT   VTETEKQWTTIHREDQVCT	ļ	1		1		1	TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL
RILRGLVSIKDKSÖQVIVQGVHELYDLEETPV SWKDDTERTNRLUTLIGRYLDKDILKQLFIAT	1		1	1	l	i	
SWKDDTERTNRLVLLGRLDKDILKQLFIAT	l	1	1			i	
VIETEKQWTTHFKEDQVCT			1	1			IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV
1999		•			1		SWKDDTERTNRLVLLGRNLDKDILKQLFIAT
FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV   GQAGLELRTSGDPPASASQSAGITGVSHLA*P   TSMPLLPFQRLCVT    GQAGLELRTSGDPPASASQSAGITGVSHLA*P   TSMPLLPFQRLCVT    FFFLRRSFAFVAQAGVWCDLGSPQPLPFGF   K*PSCLSJPSSWDYRHAPPPCPS*FLYF**RQG   FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR   CPASFYLFLKYYYLEAKFCA*GECAPSAGVGA   GYKRGHKSCLINCVVQI   GYKRGHKSCLINCVVQI   DAWGPETRLARILNPDSFIEPRPGRLPELEATR   PHMEPKASCPAAAPLMEKFHVLVGYTGSV   AALKIPLLVSKILDIPGLEVAVYTTERAKHFY   SPQDIPVTLYSDADEWEMWSRSDPVLHIDL   RRWADILVAPLDANTLGKVASGICDNLLTC   VMRAWDRSKFLICVPANNTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVCGBEGLGAMA   EVGTIVDKVKEVLFOHISGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKIGGYPHLINGGPA   LSLPRGQACSRLINVTEGPCLSFTQ*CBEAAA   GQSFIPSRASSPSCSWAQVFGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRGSGFRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA   CSTARLPH   CSTA	]	1	1	)		1	VTETEKQWTTHFKEDQVCT
FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV   GQAGLELRTSGDPPASASQSAGITGVSHLA*P   TSMPLLPFQRLCVT    GQAGLELRTSGDPPASASQSAGITGVSHLA*P   TSMPLLPFQRLCVT    FFFLRRSFAFVAQAGVWCDLGSPQPLPFGF   K*PSCLSJPSSWDYRHAPPPCPS*FLYF**RQG   FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR   CPASFYLFLKYYYLEAKFCA*GECAPSAGVGA   GYKRGHKSCLINCVVQI   GYKRGHKSCLINCVVQI   DAWGPETRLARILNPDSFIEPRPGRLPELEATR   PHMEPKASCPAAAPLMEKFHVLVGYTGSV   AALKIPLLVSKILDIPGLEVAVYTTERAKHFY   SPQDIPVTLYSDADEWEMWSRSDPVLHIDL   RRWADILVAPLDANTLGKVASGICDNLLTC   VMRAWDRSKFLICVPANNTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVCGBEGLGAMA   EVGTIVDKVKEVLFOHISGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKIGGYPHLINGGPA   LSLPRGQACSRLINVTEGPCLSFTQ*CBEAAA   GQSFIPSRASSPSCSWAQVFGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRGSGFRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA   CSTARLPH   CSTA	640	1999	A	4873	226	189	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR
GQAGLELRTSGDPPASASQSAGITGVSHLA*P   TSMPLLPPQRLCVYI   TSMPLLPPQRLCVYI   TSMPLLPPQRLCVYI   TSMPLLPPQRLCVYI   TSMPLLPPQRLCVYI   TSMPLLPPQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQPLPPGF   TSMPLLPTAQAGGVQWCDLGSPQSFLEATR   CPASFYLFLKYYLEAKFCA*GECAPSAGVGA   GYKRGHKSCLLINCVVQI   DAWGPETRLARILNPDSFIEPRGRLPELEATR   PHMEPKASCPAAPLMERKFHVLVGVTGSV   AALKLPLLVSKLLDPGLEVAVYTTERAKHFY   SPQDIPVTLYSDADEWEMKSRSDPVLHIDL   RRWADLLLVAPLDANTLGKVASGICDNLLTC   VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ   VDQLKAFGYVEPPCVAKKLVGDEGLGAMA   EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKIGGYPHLINGGPA   LSLPRGQAGSKLNWTEGPGLSFTQPCEAAA   EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKIGGYPHLINGGPA   LSLPRGQAGSKLNWTEGPGLSFTQPCEAAA   SP*LPCASNRLAFGGLIPCAPLVYPYAPFSPLL   PAPSCAPRPRAHTHISRTHPSAPLVYFAPSSRL   PAPSCAPRPRAHTHISRTHPSAPLVYFAPSSRL   PAPSCAPRPRAHTHISRTHPSAPLVYFAPSSRL   GQSPIPSRASSPSCSWAQVPGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRRGGSPRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LGSAGSGLRGBA   EVGTIVATAGGG   TSMPLTAGGG   TS	1 049	1,777	1	1 .0.5		1	FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV
TSMPLLPFQRLCVYI		ŀ	1				
	1	ł	1	1	1		
				4074	<del> </del>	127	FEEL RESEATIVACAGVOWCDLGSPOPLPPGF
FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR	650	2000	A	4874	12	437	K*ECCI CI DCCMUADHADDDCDC*EI AE**BOC
CPASFYLFLKYYLEAKFCA*GECAPSAGVGA   GYKRGHKSCLLINCVVQI	1	- }		1	İ		ETAG ADIAG NO MOTOR DICTOR DI ACC
GYKRGHKSCLLINCVVQI		İ	1	Į.		1	LIMPAKTA PUR, LUDPLINI SÓNETIKO ANTIK
A	ł	1	1	j		ŀ	
PHMEPKASCPAAAPLMERKFHVLVGVTGSV							
AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY   SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL   RRWADLLLVAPLDANTLGKVASGICDNLLTC   VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA   EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL   PAFSCAPPRAHTHSRTHPSAPLVPKPSSRAR   GQSPIPSRASSPSCSWAQVPGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRRGSGPRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA   FFFFI*DGVSLCHPGWNAVARSWLTATSASR   VQAVSCFRLPSSWDYRHATMPG*FF*YF**R   WGFTLAILVLNS*PQVICPPWPFVLTLQA   RPGIPGRFRRSWFQLP*EFPGLESLATPGD   IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW   DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP   PPCLTHLAAASCVVVWCGRWKRDSAECQCD   HSCSAVSQQEDRCRSSSCS   MNNNTTCQPSMISSMALPHYILLCIVGVFGN   TLSQWIPLTKIGKKTSTHIYLSHLVTANLLVC   G56   2006   A 4988   332   159   LVHKDMYREFFEEEAQASNKHVTRCLTSLVI   REVHIKTMR*HFLPRLEKNKNNIKD   NUSCEPEDBRAES	651	2001	A	4898	1701	771	
SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL   RRWADILLVAPIDANTLGKVASGICDNLLTC   VMRAWDRSKPILIFCPAMNTAMWEHPITAQQ   VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA   EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP   LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   LSLPRGQACSRLNWTEGPGLSFFQPGEAAA   SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL   PAFSCAPRPRAHTHSRTHFSAPLVPKPSSRAR   GQSPIPSRASSPSCSWAQVPGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRRGGFRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA   LLGSAGSGLRGEA   FFFFI*DGVSLCHPGWNAVARSWLTATSASR   VQAVSCFPLPSSWDVRHATMPG*FF*YF**R   WGFTLAILVLNS*PQVICPPWPPKVLTLQA   RPGIPGRRFRRSWFCQLP*EFEFGLESLATPGD   IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW   DPERDPGGRVQVSRGPRFGEQKTPGLQGCLP   PPCLTHLAAASCVVVWCGRWKRDSAECQCD   HSCSAVSQQEDRCRSSSCS   MNNNTTCQPSMISSMALPHYILLCIVGVFGN   TLSQWIPLTKIGKKTSTHIYLSHLVTANLLVC   G56   2006   A 4988   332   159   LVHKDMYREFFEEEAQASNKHVTRCLTSLVI   REVHIKTMR*HFLPRLEKNKNNIKD   REVHIKTMR*HFLPRL	İ	1					
RRWADLLLVAPLDANTLGKVASGICDNLLTC			İ	f			
VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA  652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKYPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPFKVLTLQA  654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKLD	1						
VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQ\$*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA FRGRQTSRPARGFSPWRPPGTMQEFSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTH_AIL_VLNS*PQVICPPWPFKVLTLQA RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPHYILLCIVGVFGN TLSQWIFLTKIGKKTSTHHYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	1			İ	1		RRWADLLLVAPLDANTLGKVASGICDNLLTC
EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA  652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTHLAILVLNS*PQVICPPWPPKVLTLQA  654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPFFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HIFLPIRLEKNKNNIKD	1		1	1	J		VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ
LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA		1	1	- [	1		VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA
LSLPRGQACSRLNWTEGPGLSFFQPGEAAA  652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPPKVLTLQA  654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPPEGLESLATPGD IPAVGLGALGVIPPVRVPQRPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HIFLPRLEKNKNNILVC		Ì		1			
LSLPRGQACSRLNWTEGPGLSFFQPGEAAA  652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPPKVLTLQA  654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPPEGLESLATPGD IPAVGLGALGVIPPVRVPQRPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HIFLPRLEKNKNNILVC		,	1	_	.	1	LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA
652 2002 A 4927 I 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTH_AIL_VLNS*PQVICPPWPPKVLTLQA  654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPRLEKKNNIKUN			1				
SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA  653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTH_AIL_VLNS*PQVICPPWPPKVLTLQA  654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPRLEKKNNIKUN	652	2002	A	4027	1	611	
PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR   GQSPIPSRASSPSCSWAQVPGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRRGSGPRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA	032	2002	1 **	1721	1	1	SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL
GQSPIPSRASSPSCSWAQVPGVALARCAGVC   KPGDSWRVAACISGRCCSRGRRRGSGPRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA	1	1					
KPGDSWRVAACISGRCCSRGRRRGSGPRNPE   QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA			1	1		1	GOSPIPSRASSPSCSWAOVPGVALARCAGVC
QSFRGAWGPSFWGSWKSQRELSAGGAQAWP   LLGSAGSGLRGEA	}		}	}	1	1	KPGDSWRVAACISGRCCSRGRRRGSGPRNPR
LLGSAGSGLRGEA	1	1	1	1			
653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW IPAVGLGALGVIPPVRVPQRPTQRSQGRGW IPAVGLGALGVIPVRVPQRWFTQRSQGRGW IPAVGLGALGVIPVRVPQRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLIPRLEKNNIKD	1	1					
VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTH.AILVLNS*PQVICPPWPPKVLTLQA  654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS  655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPRLEKKNNIKD					<del> </del>	000	PERFER CAGI CIDOUNIAMA DOME TATOA OD
WGFTHAILVLNS*PQVICPPWPPKVLTLQA	653	2003	A	4965	<sup>2</sup>	283	
654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLIPRLEKNKNNIKD	1				1		
ipavglgalgvippvrvpqrptqrsqgrgw dperdpgcrvqvsrgprfgeqktpglqgclp ppclthlaaascvvvwcgrwkrdsaecqcd hscsavsqqedrcrssscs  555 2005 A 4983 201 397 Mnnnttciqpsmissmalpiiyillcivgvfgn tlsqwifltkigkktsthiylshlvtanllvc  556 2006 A 4988 332 159 Lvhkdmyreffeeaqasnkhvtrcltslvi revhiktmr*hflprleknknnikd							
IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW   DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP   PPCLTHLAAASCVVVWCGRWKRDSAECQCD   HSCSAVSQQEDRCRSSSCS	654	2004	A	4968	3 .	437	
PPCLTHLAAASCVVVWCGRWKRDSAECQCD			i		1	1	
HSCSAVSQQEDRCRSSSCS	1	1	1	1	1	1	
HSCSAVSQQEDRCRSSSCS	Ī				J	i	PPCLTHLAAASCVVVWCGRWKRDSAECQCD
655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPHYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPRLEKNKNNIKD	1			[	1	1	HSCSAVSQQEDRCRSSSCS
TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC  656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPRLEKNKNNIKD	655	2005	- LA	4983	201	397	
656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	.   ""	2005	^	1 700		1 '	TLSOWIFLTKIGKKTSTHIYLSHLVTANLLVC
REVHIKTMR*HFLPIRLEKNKNNIKD	150	2006		1000	332	159	
ACC AND ROLL BOOK TO BE A FIGURE OF THE PARTY OF THE PART	026	2006	A	4900	332	135	
657   Z007   B   3006   123   463   WENGWICTASOD LIDGS WEEK ALL ACTION		0005	<del></del>	5000	120	165	
	657	2007	В	5008	129	403	MINOIMIE I WOOD I IDOO METELALA OFFILI EWISO

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	{	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
uence	[ ·	1	914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1.	peptide	Joquano	/=possible nucleotide deletion, \=possible
		}		sequence		nucleotide insertion
						VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
		1			ļ	AIWWEQKRQWLLQTHWTLDKYGILADARLF
658	2008	A	5017	1	292	FGPQHRPVILRLPNRRALRLX* FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
000	2008	^	3017	1	252	KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
		1	İ			HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
659	2009	A	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
				•		T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
				•	}	MAACWAVHVKTHMRPGLAVLPRLVLNSWS
		<u> </u>				*AIILLWPPKALGLQA
660	2010	Α	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
Ì		1				HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGST AGSCPROKKTTPGPTVLCVCSFWIYQRGEPH
1		· ·				HRTGARWNH
661	2011	Α	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
		1				LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
1			1			LELLGSSHPPTSASQSARITGVSHRAWPLK*F
663	2012	<u> </u>	5054	40	103	NLNQYQTLTMN
662	2012	Α	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
	ļ					LHEACLGDHVACARTLLEAGANVNAITIDGV
		l				TPLFNACSOGSPSCAELLLEYGAOAOLESCLP
		1				SPTHEGASKGHHECLDILISWGIDVDQEIPHSG
						TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
((2	2013	<u> </u>	5066	051	580	WDTPLPGAGHQSTQKLE*LFAMVEIWQ
663	2013	A	5066	951	380	VRNS*SFÄHCASVŸKHHYMDGQTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
	1			<u> </u>		GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
1		1	l			WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	A	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
		1				QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP
	1			ł		MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK
İ	1					TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC
	1					GFTVSLVTTSLVFHLPFHSSNQHE
665	2015	A	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
	1					PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE
		1				HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD
						HDQNEGFHCREECRILGHSDRCWMPRNPMPI
)						RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
	1					ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
		1				PTKPSVSYEIVDPGITARRC_
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
<u></u>	1	<u> </u>	<u> </u>			VGPKAQLSGLQLQPCLYKRR
667	2017	A	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
668	2018	A	5086	852	233	PCVY*RR NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
""	2018	^	3000	0.72	233	RALPTTFADIENLKYLLFTRDASQPFYLGHTV
						IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
<b>j</b>	]	]				TCADQSVIWKLSEDKQLAICLKYAGVHAENA
		1				EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
						CCSDMAITFNGLTPQKMEVMMYGLYRLRAF
669	2010	<del>   </del>	1	<del> </del>	329	GHYFNDTLVFLPPVGSEND
1 009	2019	A	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
		1	1			RGCQHEAAPCPRGPGSDGLHHASAACASLPP
						SPILPVLLPELGPL
670	2020	A	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP
						······································

WO 01/57188

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D-Aspartic Acid, E-Glutamic Acid, F-Phenylalanine, G-Glycine, H-Histidine,
nucl-	peptide	]	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino	Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
		1		amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	residue of	sequence	Y=1 yrosine, X=Unknown,Stop codon,
		ļ		peptide	]	/=possible nucleotide deletion, \=possible
	ĺ	i	i	sequence		nucleotide insertion
						DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI
	1	ļ		1		VGGTDQQYVSNDSGIYVSRIKENGAAALDGR
	1				]	LQEGDKILSVNGQDLKNLLHQDAVDLFRNA
	i	ì				GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
	1			·		FMVLVPVFALTMVAAWAFMRYRQQL
671	2021	Α	5105	672	400	RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF
		1	1			VLLLLLISLLCLYWKARKLSTLRSNTRKEKA
		1	<b>j</b>		İ	LWVDLKEAGGVTTNRMED*EEDECN
672	2022	A	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT
0.2	2022	1			}	YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS
	1	1	1	1		NOAHGALOEYVLAPCS
673	2023	A	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA
015	2025	^	1 222	1		NHFVEVT
674	2024	A	5153	3	2953	LTEDOPFDILOKSLOEANITEQTLAEEAYLDA
0/4	2024	Ι ^	3133	"	1 2733	SIGSSOOFAOAOLHPSSSASFTQASNVSNYSG
		.]	1	1		QTLQPIGVTHVPVGASFASNTVGVQHGFMQH
	ł	1	i		ĺ	VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS
1	į.		ł	1		MMTINNLDGSQIILKGSGQQAPSNVSGGLLV
	ŀ		-	1		HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF
Ì	1.	1	1	1	ļ	QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIQM
ļ	Í	1	1		1	GQQNTYNVNNLGIQQHHVQQGISFASASSPQ
	1	1	ĺ	Į.	1	GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG
		}	1	Ì	ļ	GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH
	İ	1		Į.	Į.	HVQTINGQLLQTQPSQLISGQVASEHVMLNR
	1	1				NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ
ł	1	1	1	1	Ī	TFAASGSPVIANHASPQLVGGQMPLQQASPT
ĺ	1		1		l .	VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR
1		1	l		1	FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL
		1	1	ļ	}	TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT
1	1	1	1			GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR
ļ		-	-		1	QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ
1	0	-				QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ
İ	ŀ	1	1	İ	i	QEKY VOSSPUMPA V V ESMSOUQAM AAAQ
1	• ]	1	ļ	1		LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD
						AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV
	1	1		1		ATOLLKRTQAMLNKYRCLLLEDAMRINPPAE
	1.		1	1	1	MVMIDRMFNQEERASLSRDKRLALVDPEGFQ
İ					1	ADFCCSFKLDKAAHETQFGRSDQHGSKASSS
	1	-			1	LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP
]	]	1				NHIVVSAEGNISKKTECLGRALKFDKVGLVQ
		1		l		YQSTSEEKASRREPLKASQCSPGPEGHRKTSS
1		1		1		RSDHGTESKLSSILADSHLEMTCNNSFQDKSL
İ		1	1	1		RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET
			1	Į	1	TFKNILELKKAGRQPQSDPTVSGSVELDFPNF
1	1			1	1	SPMASQENCLEKFIPDHSEGVVETDSILEAAV
1						NSILEC
675	2025	A	5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR
1 3,3	1 2023	1.,		1		LYDEVQEVVYFPAVVHDNLGERLKCTYIEID
		ĺ		1		QVPETYAVVLSRPAWLWGAEMGANEHGVCI
		1				GNEAVWGREEVCDEEALLGMDLVRLGLERA
		1.		1 '		DTAEKALNVIVDLLEKYGQGGNCTEGRMVF
1		1				SYHNSFLIADRNEAWILETAGKYWAAEKVQE
1	4	- {			1	GVRNISNQLSITTKIAREHPDMRNYAKRKGW
1			ì	1	1	WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE
1		-	1	Į.	1	GYKLLNKHKGNITFETMMEILRDKPSGINME
1	1	1	1		1	GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER
1	- [	- 1	ļ	1	}	SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS
1		ı	1	(	[	HFKPDRRHPLYQKHQQALEVVNNNEEKAKI
1	l	1	l l	- 1		MLDNMRKLEKELFREMESILQNKHLDVEKIV
}	- 1	İ	ł			NLFPQCTKDEIQIYQSNLSVKVSS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
676	2026	A	5155	sequence 2	306	nucleotide insertion  FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG FTLLARMVSIS*PHDPPASASQSAGITGVSHRA
677	2027	A	5167	97	740	RPT FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR LNKRSFFMISPTDQQVHCWAWLKKHMPKDS NLLLEDVTWKYTALNLIGPRAVDVLSELSYA PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT GEPGFMLYIPIEYRWGFTMLSTLVSNS
678	2028	A	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRQ GRIAKMPVKWIAIESLADRVYTSKSDVWAFG VTMWEIATRGMTPYPGVQNHEMYDYLLHG HRLKQPEDCLDELCKI**SPQSP
679	2029	A	5190	39	499	RESOVKHFKMRKIDLCLSSEGSEVILATSSDE KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVSAEGTVVSNLSS
680	2030	A	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG
681	2031	A	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS
683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK OSESAI
684	2034	A	5220	1	194	NLMKEMONLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS
685	2035	Α -	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	A	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT
688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide scq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FSSTMSLAKLLQERGISAKVYHSPISENPLQPL PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN FLASRPAETFLQEMYGLRPSRNPPDVGQLKM NLVDRLKRLGIARVVKNPGAQENGRCQEAEI GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM GSFAAPVCTSSPKMGVLKED
689	2039	A	5254	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS GAPAGARGGPAKANSNPFEVKVNRQKFQILG RKTRHDVGLPGVSRARALRKRTQTLLKEYKE RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA LEQQRHHEKKSIYNLNEDEELTHYGQSLADIE KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG GLLHKKTQQEGEEREKPKSRKELIEELIAKSK QEKRERQAQREDALELTEKLDQDWKEIQTLL SHKTPKSENRDKKEKPKPDAYDMMVRELGF EMKAQPSNRMKTEAELAKEEQEHLRKLEAE RLRRMLGKDEDENVKKPKHMSADDLNDGFV LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA SDPESNEEEGDSSGGEDTEESDSPDSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMEEQ LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK
690	2040	A	5261	1	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEJFPILIVLKFYLVTLT SFVK
691	2041	A	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV
692	2042	A	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK
693	2043	A	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA
694	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA
695	2045	A	5315	125	1596	ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

	1 200 TD	N/a4	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide			location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-		USSN	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496		acid residue	O=Glutamine, R=Arginine, S=Serine,
uence		İ	914	ng to first		T=Threonine, V=Valine, W=Tryptophan,
	1			amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
	Į.	ł	! '	residue of	sequence	/=possible nucleotide deletion, \=possible
		ļ	l	peptide	1	
	1	i	ł	sequence		nucleotide insertion
					1	TGSGKEEGPAPCKQMKLEAAGGPSALNFDSP
	İ		i	1		SSLFESLISPIKTETFFKEFWEQKPLLIQRDDPA
		i	ì	ļ		LATYYGSLFKLTDLKSLCSRGMYYGRDVNV
	l	1		1	1	CRCVNGKKKVLNKDGKAHFLQLRKDFDQKR
		ł	ļ			ATIQFHQPQRFKDELWRIQEKLECYFGSLVGS
		ļ	1		1	NVYTTPAGSOGLPPHYDDVEVFILQLEGEKH
	1		1			WRLYHPTVPLAREYSVEAEERIGRPVHEFML
	ł		f .			KPGDLLYFPRGTIHQADTPAGLAHSTHVTIST
	i	1	1			YONNSWGDFLLDTISGLVFDTAKEDVELRTG
1	ì	ĺ	1	1		IPRQLLLQVESTTVATRRLSGFLRTLADRLEG
1	1	Į.	ı	1		TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP
	1			İ		GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD
i		1		i	,	ETQEKMVYIYHSLKNSRETHMMGNEEETEFH
	1	1	{	{		GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE
ļ		1	1			
]						EKESLVLSLWTECLIQVV
696	2046	A	5318	1476	742	LMKXYLEAAELGEISDIHTKLLRLSSSQGTIET
į	ł	1				SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN
1	Į.		l	ļ	•	VEKLQVLLNCMTEIYYQFKKDKAERRLAYN
i	1	1	ſ		1	EEQIHKFDKQKLYYHATKAMTHFTDECVKK
	i	]	1			YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI
Ì	ŀ	İ		1		EEEVSKYQEYTNELQETLPQKMFTASSGIKHT
1			ì	1		MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE
1	l l	1	ļ			LAENNHILESGGSLTMDGGLRNVDCL
607	2047	A	5320	244	478	LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP
697	2047	^	3320	277	1770	PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG
1	ı	}	1	1		VSPSWPGWSRTPDFR
		+	+	266	714	LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC
698	2048	Α	5324	200	/14	LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP
1	l	1				VGGLLMAFQKYSGETVQERKQKDRKALHEL
1	1		ì			KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA
	<b>!</b>		1			KLEEWKOKLQV IEHLYENIESSLQEDEFENDA
1	·	1				KKIEALLNLPRNPSVIDKQDKD
699	2049	Α	5334	699	277	RPHGHLVCISSSAGLSGVNGLADYCASKFAA
	1		1		İ	FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM
ı		1	1	1	Ì	FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM
			1		Į	YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI
	1			1		LHAMDGFADQKK
700	2050	A	5344	3	614	PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE
1,00	2000		-5	1	1	OIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS
1				1	1	OTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
1				ı	i	PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
				1		VETIOAOLLSTHDQPSVQALADEKNGAQTRP
						AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK
				1 .		LLATNGTPL
<u></u>		<del></del>	<del></del>	<del></del>	1202	HASVLFCRVMAASKTQGAVARMQEDRDGSC
701	2051	A	5346	3	1383	STVGGVGYGDSKDCILEPLSLPESPGGTTILE
		1	1	l	1	
	1	1	l l	1		GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV
ŀ		1	1	1		IADVKL VADFQRYIL YWRKRFTEQPITDFCSV
1		-		1		IRINSTAPFEEQENYFLLCDVLPEDRILREELQ
		1		1	1	KQRLREILEQQQQERNDTNFHGVCMFCNEEF
1		1		ſ		LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL
				1		CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR
		1	1	1		KKQHRKINPKNREYDRFYVINYLELGKSWEE
	1	ļ	1	1		VOLEDDRELLDHQEDDWSDWEEHPASAVCL
1	İ	l l	1	1	1	FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG
		1			1	LNFYOOVKLVNFIRRQVHQCRCYGCHVKFKS
1	l l	ı	1			KADLRTHMEETKHTSLLPDRKTWDQLEYYFP
1					1	
1			1	1	ì	TYENDTLLWTLSDSESDLTAOFONENVPIISE
1		Ì				TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKOSSILNOLLL
702	2052	A	5356	2502	1540	TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL MAAATRGCRPWGSLLGLUSAAAAAWD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LASLRCTLGAFCECDFRPDLPGLECDLAQHL AGQHLAKALVVKALKAFVRDPAPTKPLVLSL HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILLQELEPVISRAVLDNPHHGFSNSGI MEERLLDAVVPFLPLQRHHVRHCVLNELAQL GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK TVASRIAFFL
703	2053	A	5380	278	657	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK KVKQDLLELKSVIKLQAWWRGTMIRREIGGF KM
704	2054	A	5381	1	1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV DMSDLSPEEQWRVEHARMHAKHRGHEAMH AEMVLILIATLVVAQLLLVQWKQRHPRSYN MVTLFQMWVVPLYFTVKLHWWRFLVIWILF SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA MDFGISLLFYGLYYGVLERDFAEMCADYMA STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK QTCFYCKEKVDLKRMFSNPWERPHVMYGQL LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP GNKRPLSALYRLESKEPFLSVGGYVFDYDYY RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI S
707	2057	A	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	A	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	A	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK GKWFCPRCVQEKRKKK
710	2060	A	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY EESDSDESWTTESAISSEAILSSMCMNGGEEK PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV SAEIIRKMQQ
711	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD LYRDIPELQGF
712	2062	A	5499	91	749 -	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KAPELLQGQSEDEQPDASQMHVYSLGMTLY WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH
713	2063	A	5506	22	478	RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR RLVGLVLGTISEVSREPCFSSSCWSCVAIKI VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT
/13	2003	A	2300		478	TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL GGVECLLLAVMAYDRCVAICKPLHYMVIMN PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR CGHHEVDHFLCEMPALIRMACISTV
714	2064	A	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR RLTSSVSCALDEAAAALTRMRAESTANAGQS DK
715	2065	A	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM GRTALFHHSGGSSGYESLRRDSEATGSASSAP DSMSESGAASPGARTRSLKSPKKRATGLQRR RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG LACVSTRLRLAERRQQRLREVQAKHKHLCEE LAETQGRLMLEPGRWLEQFEVDPELEPESAE YLAALERATAALEQCVNLCKAHVMMVTCFD ISVAASAAIPGPQEVDV
716	2066	A	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD SSSSERTMDLVLEMCNTNSIHWCGISGRQLG KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	A	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAQENA
719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC
720	2070	. ·	5628	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG
721	2071	A	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN
722	2072	A	5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

						Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	l	09/496	correspondi	to last amino	O=Glutamine, R=Arginine, S=Serine,
uence	1	ļ	914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
donos		1	i	amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	ł	l .	residue of	sequence	/=possible nucleotide deletion, \=possible
1	}	1	1	peptide	}	/=possible nucleonide deterion, /-possible
1			Ì	sequence		nucleotide insertion
	<del></del>	<del>                                     </del>	1			QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL
ļ.	l .		1	1	1	GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
<b>ļ</b>		1.	1	1		AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI
1	٠ ،	ł	1	1	1	PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN
1	1	ì	1	1		KAEINPRINGDRAACQIRILRFVSNVTEADRI
1	1	1		1		LLRWERYEPLEARDLLSFIVYYKESPFQNATE
1	1	1	1	ł		HVGPDACGTQSWNLLDVELPLSRTQEPGVTL
	}	1	1	ł	1	ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS
İ	}	1	ŀ		7	PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW
1			1	1		KPPTQRNGNLTYYLVLWQRLAEDGDLYLND
1		1	1		1	YCHRGLRLPTSNNDPRFDGEDGDPEABMESD
	1	1	1 .	1	1.	CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH
1	1	1		1	1	NATTIPISPWKVTSINKSPORDSGRHRRAAGPL
İ		1	.]	Į.	1	RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY
	į			ł	1	RIDIHACNHAAHTVGCSAATFVFARTMPHRE
1	1	1		1 .	1	ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL
}	1	ł	1	i		ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV
		- [				HLALLPPGNYSARVRATSLAGNGSWTDSVAF
	1	1		1	1	YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA
	ì	1		1		LGFFYGKKRNRTLYASVNPEYFSASDMYVPD
		1				EWEVPREQISIIRELGQGSFGMVYEGLARGLE
1			1	Ļ		AGEESTPVALKTVNELASPRECIEFLKEASVM
}		1	<b>\</b>	<b>\</b>	l	KAFKCHHVVRLLGVVSQGQPTLVIMELMTR
		1	1	l	· ·	GDLKSHLRSLRPEAENNPGLPQPALGEMIQM
ł		ì				AGEIADGMAYLAANKFVHRDLAARNCMVSQ
			1	1		DFTVKIGDFGMTRDVYETDYYRKGGKGLLP
1	1		1	Į.		VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV
	1	1	ŀ		1	TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
l l	1	}	1	1	1	PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL
		.	1	Į.		RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP
				<u> </u>	ŀ	TPRICSPONGGPGH
	2073	-	5672	1	216	LAWIDNILPEKEKKETDKKRKKKGAHEDCD
723	2073	A	3072	1.		EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR
		- 1		l	1	SSSWSVSI.
			5704	4235	940	ARGRESRPVWAASWGGRGRPAARRRPRGLA
724	2074	A	3704	4233	1770	ATMGFEL DRFDGDVDPDLKCALCHKVLEDP
		1		1	1	1.TTPCGHVFCAGCVLPWVVQEGSCPARCRGR
i	1				+	I SAKELNHVLPLKRLILKLDIKCAYATRGCGR
i	1	l			1	VVKLOOLPEHLERCDFAPARCRHAGCGQVLL
	1	- 1	1			PROVEAHMRDACDARPVGRCOEGCGLPLIH
	i	1		l l		GEORAGGHCCARALRAHNGALQARLGALHK
		. [	1	1	1	ALKKEALRAGKREKSLVAOLAAAQLELQMT
Ì	ŀ	-	i		l	ALRYOKKETEYSARLDSLSRCVAAPPGGKGE
1	ł	1	1	1	1	ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG
		- 1	1	l l		SSSEGIEVSKIVDSGPAAKEGGLOIHDRIIEVN
		-	1		1	GRDLSRATHDOAVEAFKTAKEPIVVQVLRRT
		1				PRIKMETPPSESOLVDTGTOTDITFEHIMALI
			1	1		KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI
1	- 1	- 1	1		l	GDIHOEMDREELELEEVDLYRMNSQDKLGLT
1	1	1	1	1		VCVRTDDEDDIGIYISEIDPNSIAAKDGRIREG
1	ļ	i i	1	Į	ł	DRIIQINGIEVQNREEAVALLTSEENKNFSLLI
		1		l l	1	ARAELQLDEGWMDDDRNDFLDDLHMDMLE
1	1	- 1		[	{	EQHHQAMQFTASVLQQKKHDEDGGTTDTAT
1	}	İ	1	1	1	ILSNQHEKDSGVGRTDESTRNDESSEQENNG
1	1	-			1	DDATASSNPLAGQRKLTCSQDTLGSGDLPFS
1	1	1		i i	1	NESFISADCTDADYLGIPVDECERFRELLELK
ı	t				1	NESFISADCTDADYLGIPVDECERTREDELER CQVKSATPYGLYYPSGPLDAGKSDPESVDKE
	i		1	1 -	. 1	I COVESATE IOUT I LOGGEDIAGEOUS TO A DECE
			1	1	1	THE VALUE DOING BOT CHAD VARAMOUT KEUAL
						TELLNEEL RSIELECL STVRAHKMOOLKEQYK
	,					LELLNEELRSIELECLSIVRAHKMQQLKEQYR ESWMLHNSGFRNYNTSIDVRRHELSDITELPE KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRE

NO. of not peptide estide sequence peptide estide sequence peptide estide sequence where the contemporary process and sequence where the contemporary process and sequence peptide sequence pepti	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide of the state of the							D=Aspartic Acid, E=Glutamic Acid,
Sequence   Sequence			1100				F=Phenylalanine, G=Glycine, H=Histidine.
uence uence   09/496   1914   ng to first amino ecid residue of peptide residue of peptide sequence   1914   ng to first amino ecid residue of peptide sequence   1914   ng to first amino ecid residue of peptide sequence   1915   ng to first amino ecid residue of peptide sequence   1915   ng to first amino ecid residue of peptide sequence   1915   ng to first amino ecid residue of peptide sequence   ng to first amino ecid residue   ng to first amino ecid residue   ng to first amino ecid residue   ng to first amino ecid residue   ng to first amino ecid residue   ng to first amino ecid residue   ng to first amino ecid residue   ng to first amino ecid residue   ng			l				
### Section of first amino acid residue of peptide sequence peptide sequenc			1				M=Methionine, N=Asparagine, P=Proline,
amino acid residue of peptide requence peptides of peptide sequence peptides sequence peptides peptides peptides sequence peptides peptide	•	licities .	1				
Peptide   Pept	ucticc			714			T=Threonine, V=Valine, W=Tryptophan,
peptide   peptide   peptide   peptide   pelebidide insection   possible   pelebidide insection   pelebidide				}			V=Tyrosine, X=Unknown, *=Stop codon.
						30quenoe	
AAGGICPSSEGAVGTTEAYGPASNLT						ļ	
DPEVGTPTYSSILEIDPROPLESKERRAS   SRSPTPSOKLGAYLPSYHASPYKHAHPIA   QHYOSYMQLIQQXSAYEYAQSQMSLYS)   DLSSPTPSBERMEWKYKIRSDGTRYTTKRE    DRILRRALKIRERSGMTIDDIAVSEME    GRYWSKEERKQHLVKAKEQRRRERFIN    LOLKEQQAADDRKEMNILBLSHKKMM    LOLKEQQAADDRKEMNILBLSHKKMM    LOLKEQQAADDRKEMNILBLSHKKMM    LOLKEQQAADDRKEMNILBLSHKKMM    LOKKIPDNWMTQELITHGTKSPDGTRVY    LSVTTV				<del> </del>	Scyttoffee		
SRSPIPSQKLGSAYLESYHHSPYKHHSPYKHALSOKSMSLVSM   OHYOSYMQLOQKSAVEVASQSMSLVSM  DLSSPIPSEPRMEWKYKIRSDGTRYTIKER   DRILIRERALKIREERSGMTIDDDAVSEME   GRYWSKEERSCHULVAKEQRERRERMM   LDCLKEQQAADDRKEMNILEISHKKMN   NKKIPDNWMTOPLITHIGTSEPDGTRVY    LSVITTV		1	1		[		
OHYOSYMOLLOQUSAVEYAGSOMSLUS   DISPPTSEERMWW XURSDOTRNYTIKR    DRILRERALKIREERSGMTTDDDAVSEME    GRYWSKEERKQHL VKAKEQRRRERING    LOCKEQQAADDREMMILEISHKKMM    LOCKEQQAADDREMMILEISHKKMM    NKKIPDNWMTIQELLTHGTKSPDGTRVY    LSVITIV     LSVITIV     SVITIV     SVITIV     GISTEVSEAPVANDKPKILVVKVQKKKKNAV     LCGKNGGGAFLIPYFLTLIFAGVPLFLLEC     QYTSIGGLOVWKLAPMFKOVLGAAVLIE     LNIYIVIISWAMYYLYNSTITLIPKQCD     WINTDR.CFSNYSMYNTINMTSAVVEEW    MHQMTDGLIDKPGQRWPHALTIAAWUL     CIWGGVGWTGKVVYFSATYPYIMLIIFT     TLOCARGGLIPTIPFRYALSDES WULDA     IFFSYGLGLGSLIALGSYNSPHNIVYKSI     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYPRLIFT     LLML.GIDSCFOTVEGFTIALVDEYVELD     ASGMSLIFLVFTECVSISWFYGVNRFYTCH     ASGMSLIFLVFTECVSISWFYGVNRFYTCH     ASGMSLIFLVFTECVSISWFYGVNRFYTCH     LLML.GIDSCFOTVEGFTIALVDEYVELD     ASGMSLIFLVFTECVSISWFYGVNRFYTCH     ASSMSLIFLVFTECVSISWFYGVNRFYTCH     ASSMSLIFLVFTECVSISWFYGVNRFYTCH     ASSMSLIFLVFTECVSISWFYGVNRFYTCH     TLMARKONSOFT     T			ł			ļ	
DIASPTYSEPRMEWKYKIRSDGTRYITED   DIASPTYSEPRMEWKYKIRSDGTRYITED   DIASPTYSEPRMEWKYKIRSDGTRYITED   DIASPTYSEPRMEWKYKIRSDGTRYITED   DIASPTYSEPRMEWKYKIRSDGTRYITED   DIASPTYSEPRMEWKYKIRSDGTRYITEN   DIASPTYSEPRMEWKYKIRSDGTRYITEN   DIASPTYSEPRMEWKYKIRSDGTRYITEN   DIASPTYSEPRMEWKYKIRSDGTRYITEN   DIASPTYSEPRMEWKYKIRSDGTRYITEN   DIASPTYSEPRMEWKYKIRSDGTRYITEN   LINTYIVISWALIPLAYOPEPILLSC   QYTSIGGLOVWKLAPMFKGVGLAAAVL   LINTYYIVISWALIPLAYOPEPILLSC   QYTSIGGLOVWKLAPMFKGVGLAAAVL   LINTYYIVISWALIPLAYOPEPILLSC   QYTSIGGLOVWKLAPMFKGVGLAAAVL   LINTYYIVISWALIPLAYOPEPILL			1			1	
DRILIRERALKIREERSGMITDDDANSEMI					•	ļ	
GRYWSKEERKORILVKAKEORRREFMM   LDCLKEQQAADDRKEMNILEISHKKMM   NKKIFDNWMTIQELLTHGTKSPDGTRVY  LSVTTV   L	ļ	}	j	ļ	1	}	
LDCLKEQQAADDRKEMNILĒLISIKKKMM NKKIPDNWMTIQELLTHGTKSPDGTRVY LSVTIV		ļ	1		]	1	
NKKIPDNWMTQELLTHGTKSPDGTRVYY		Ì					
1.58\text{TIV}		1	1	1			
725   2075   A   5707   3   1770   QISTEVSEAPVANDKPKTLVVKVQKKAAD   DRDTWKGRFDFLMSCVGYAIGLGNVWR   LCKNGGGATLIPYLTLIFAGVPLFLLEC   QYTSIGGLGVWKLAPMKGYGLAAVLS   LNIYYYIISMATIYY INSPTITLPWKQCD   WNTDRCFSNYSMVNTTNNTSAVVEFWE   MHQMTDGLDKPGQIRWPLAITLAIAWILL   CIWKGYGWTGKVVYFSATYPYIMLILFF   TLPGAKEGILFYITPNFRKLSDSEVULDA   IFFSYGLGGLGSLIALGSYNSFHNNVYRDSI   CINSCTSMFAGFVIFSIVGFMAHVYKRSIA   AASGPGLAFLAYPEAVTQLPISPLWAILFF   LLMLGIDSQFCTVEGFTIALVDEYPRLKR   ELFIAAVCIISYLIGLSNITIQGGIYVFKLFD   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFFECVSISWFYGVNRPYCHV   ASGMSLFLVFECVSISWFYGVNRPYCHV   ASSMSLFLVFFECVSISWFYGVNRPYCHF   ASSMSLFLVFECVSISWFYGVNRPYCH   ASSMSLFLVFECVSISWFYGVNRPYCRF   ASMMSLFLVFECVSISWFYGVNRPYCRG   ASSMSLFLVFECVSISWFYGVNRPYCRG   ASSMSLFLVFECVSISWFYGVNRPYCRG   ASSMSLFLVFECVSISWFYGVNRPYCRG   ASMMSLFLVFECVSISWFYGVNRPYCRG   ASSMSLFLVFECVSISWFYGVNRPYCRG   ASSMSLFLVFECVSISWFYGVNRPYCRG   ASSMSLFLVFECVSISWFYGVNRPYCRG   AMSKNSSQETCTRAFILAQGGSKH   AWFEKMTCYLQLFNICLPDVSEE   IQASRASPYPRVKVDFALSCHEDLLAPISI   WKYHSPEEEISLGPACWLWDFLRSQQA   LPLSGGVDSAATACHVSDFLRSQQA   LPLSGATSAATAPOGRAGAAGGSKRAAA   LEPLADQVSQTDEEDMGMTYAELSVY   RVAKMGPYSMFCLIGNVCSSADIN   SKTDLRAFACGSSCBATACHVSDFLRSQQA   ASSMSLRAFAPOGRAGAGKRACHAD   DARRIKARAFAPOGRAGAGKRACHAD   ASSMSLRAFAPOGRAGAGKRACHAD   AGTVYLAAVNRLYQLSGANLSLEAEA   AAGTVYLAAVNRLYQLSGANLSLE	Ì	1		i			1
DRDTWKGREPDLMSCVGYLGLGNVWR			<del>  </del>		<u> </u>	1000	
LCGKNGGGAFLPYFLTLIFAGYPLFLLEG   QYTSIGGLGWWKLAPMFKGVGLAAAVLS   CHITYYIVIISWAITYLYNSFITTLPWKQCD     WNTDRCFSNYSMVNTTIMTSAVVEWE     WHOMTOGLDKPGQRWPLAITLAIAWILY     CIWKGVGWTGKVVYFSATYPYIMLIILFF     TLPGAKEGILFYITPNFRKLSDSEVWLDA.     IFFSYGLGLGSLALGSYNSFHINNVYRDSI     CINSCTSMFAGFVIFSIVGFMAHVTKRSIS     AASOPGLAFLAYPEAVTOLPISPLWAILFF     LLMLGIDSQFCTVEGFITALVDEYPRLLRI     ELFIAAVCLISYLIGLSNITGGIYVFKLFD     ASGMSLLFLVFECVSISWFYGVNRFYDN     MYGSRPCIWKLCWSFFTPIVAGVFIEN     ANGBALLFLVFECVSISWFYGVNRFYDN     MYGSRPCIWKLCWSFFTPIVAGVFIEN     AMFULT MAY SOPPOLA     ANTHER CONTRACT     ANTH	725	2075	A	5707	3	1770	DEDLINK OBEDEL WOONGANIGI CHIMIDEDA
QYTSIGGLGWKLAPMFKGVGLAAAVLS   LNIYYIVIISWAIYYLYNSFTTTLPWKQCD   WNTDRCFSNYSMYNTTNMTSAVEFWE   MHQMTDGLDKRGQRWFLATLAIAWILL   CIWKGVGWTGKVVYFSATYPYIMILIFF   TLPGAKEGILFYITPNFRKLSDSEVWLDA.   IFFSYGLGLGSLIALGSYNSFHNNYYRDS    CINSCTSMFAGFYFSIVGFMAHVTKRSIA   AASOPGLAFLAYPEAVTQLPISPLWAILFF   LLMLGIDSQFCTVEGFITALVDEYPRLRI   BLFIAAVCIISYLIGISNITQGGIVYFKLFD    ASGMSLIFLVFFECVSISWFYGVNRFYDN    MYGSSRCIWWKLCWSFFTPIIVAGVIFSP,   MYGSSRCIWWKLCWSFFTPIIVAGVIFSP,   MYGSRPCIWWKLCWSFFTPIIVAGVIFSP,   MYGSRPCIWWKLCWSFFTPIIVAGVIFSP,   MYGSRPCIWKLGWSFFTPIIVAGVIFSP,   MYGSRPCIWKGLGWSFTPIIVAGVIFSP,   MYGNAYMFLTLKGSLKQRQVMVQPSE     PRRDPGRTPELRGSAPRKTGANMPVRRG     PQNTFLGTIIRKFEQQNKKFIANARVQNC     YCNDGFCEMTGSRPDVMQKPCTCD     HASEYFFKLCSFQVFLSFPLATIVIDVGLV     LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI     AWFEKMTCYLQLFNICLPDVSEE     T28   2078   A 5737   1899   649   IQASRASPYRVKVDFALSCHEDLLAPIS     WKYHSPEETSLGACWLWDFLRSQQA     LPLSGGVDSAATACLIYSMCCQVCEAVR     EEVLADVRTIVNQISYTPOPROLCGRL     YMASSNSSGETCTRARELAQQIGSHHIS     PAVKAVMGJPSSLGYLTKYDGSSADINI     ISKTDLRAFVOFCQRFQLSHGSHHIS     PAVKAVMGJPSSLGYLTKYDGSSADINI     ISKTDLRAFVOFCQRFQLSHGSHHIS     PAVKAVMGJPSSLGYLTKYDGSSADINI     ISKTDLRAFVOFCQRFQLSHGSHHIS     PAVKAVMGJPSSMFCKLLGMWRHICTPRQ     KVKRFFSKYSMRRKMTTLTPAYHAEN     KVKRFFSKYSMRRKMTTLTPAYHAEN     DNRFPLRPFLYNFYALSSYN     RKVAKMGPYSMFCKLLGMWRHICTPRQ     KVKRFFSKYSMRRKMTTLTPAYHAEN     DNRFPLRPFLYNFALAGAGAGRKRLAD     DNRFPLRPFLYNFALAGAGAGRKRLAD     DNRFPLRPFLYNFALAGAGAGRKRLAD     PASRRLRAFGSRRRLAPGTRRAAQPAHA     PRAAGGAPLSARAAAASPPFOTPRCPV     LLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTPRCPV     LLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTPRCPV     LLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTPRCPV     LLLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTPRCPV     LLLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTPRCPV     LLLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTPRCPV     LLLLLGAARAGALEIQRRFPSTPTNNFA     AAGGAPLSARAAAASPPFOTRRAAQPAHA     PRAAGGAPLSARAAAASPPFOTRRAAQPAHA     PRAAGGAPLSARAAAASPPFOTRRAAQP	{ ·	1	1	1 .			TOOMIGGGAET IDVELTE IEVONDI ELLE EGGLO
LINYYIVIISWAIYYL YNSETTITLPWKQD						I	OVICE CAMANA VENERAL VENERAL OFFICE COPING CONTROL OF THE PERSON OF THE
WNTDRCESNYSMVNTTINMTSAVVEFWE   MHQMTDGLDKPQGIRWPLAITLAIAWILN   CIWKGVGWTGKVVYFSATYPYIMLIILFF   TLPGAKEGILFYITPNFRKLSDSEVWLDA   IFFSYGLGLGSLIALGSYNSPHINNYRDSI   CINSCTSMFAGFVIFSIVGFMAHVTKRSIA   AASGPGLAFLAYPEAVTQLPISPLWAILFF   LLMLGIDSQCTVEGFITALVDEYPRILRS   LLMLGIDSQCTVEGFITALVDEYPRILRS   MYGSRPCIWWKLCWSFFTPIIVAGVIFISX   MYGSRPCIWWKLCWSFFTPIIVAGVIFISX   MYGSRPCIWWKLCWSFFTPIIVAGVIFISX   MYGSRPCIWWKLCWSFFTPIIVAGVIFISX   MYGRPCIWWKLCWSFFTPIIVAGVIFISX   MYGRPCIWGKLGWSFYGANMPVRG   RPENGPEQPQAGSSTSKEAY    PRRDPGRTPELRGSAPRKTGANMPVRRG   PQNTFLGTIRKFEGQNKKFILANARVQNC   YCNDGFCEMTGFSRPDVMQKPCTCD   HASEYFFKLCSFQVFLSFPLATIVOGIV   LVKSPNVHYVYLLLVLSGLFPYPLHFF   AVFEKMTCYLQLLFNICLPDVSEE   A 5737   1899   649   IQASRASPYPRVKUDFALSCHEDLLAPISI   WKYISPEESISJOPACWLWDFLRRSQQA   LPLSGGVDSAATACLIYSMCCQVCEAVR   EEVLADVRITVNQISYTPQDPRDLCGRIL   YMASKNSSQSCTCTRARELAQQIGSHHISL   PAVKAVMGIFSLVTIKSPLFAAHGGSSR   ALQNVQARIRMVLAYLFAQLSUCARD   LLVLGSANVDESLLGYLTKYDCSSADINI   ISKTDLRAFVQFCIQRFQLPALQSILLAPA   LEPLADGQVSDEEDIMGMTYAELSVY   RKVAKMGPYSMFCKLLGMWRHICTPRQ   KVKRFFSKYSMNRHKMTILTPAYHAEN   DNRPDLRPFLYNTSWPPWGFRCIENQVLQ   AEPQSLDGVD   PGCAARLSRARAPGFGAAGAGRKRLAD   PASRRILRAPGSRRAAACASPPPFQTPPRCP   LLLLGAARAGALEIQRRFPSPTPTNNFA   AAGTVYLAAVNRLYQLSGANLSLEAAA   ARGTYYLAAVNRLYQLSGANLSLEAAA   ARGTYYLAAVNRLYQLSGANLSLEAAA   AAGTVYLAAVNRLYQLSGANLSLEAEA   AAGTVYLAAVNRLY	1	[	1	1			
MHOMTOGLDKPGQIRWPLATILAIAWILA   CIWKGVGWTGKVVYFSATYPYIMLIILFF   TLPGAKEGILFYITPNFRKLSDSEVWLDA.   IFFSQLGLGSLIALGSYNSFHNNVYRDSI   CINSCTSMFAGFVIFSIVGFMAHVTKRSIA   AASGPGLAFLAYPEAVTQLPISPLWAILFF   LLMLGIDSQFCTVEGFTALVDEYPRLKRIE   LEFLAVCIISYLIGLSNITQGGIVYFKLFD   ASGMSLLFLVFECVSISWFYGVNRFYDN   MVGSRPCIWWKLCWSFTFIIVAGVFIFS.   MTPLTMGNYVFKWGQGVGWLMALSS   IPGYMAYMFLTLKGSLKQRIQVMVQPSE   RPENGPGQPQAGSSTSKAYI   PRRDPGRTEELRGSAPRKTGANMPVRRG   PQNTFLGTIIRKFEQQNKKFILANARVQNC   YCNDGFCEMTGFSRPDVMQKPCTCD   HASEYFFKLGSFQVFLSFPLATIVIDVGLV   LVKSPNVHYVYVLLLVLSGLLFYPLIHFI   AWFEKMTCYLQLLFNICLPDVSEE   A 5737   1899   649   IQASRASPYPKVDFALSCHEDLLAPISI   WKYHSPEEEISLGPACWLWDFLRRSQQA   LPLSGGVDSAATACLIYSMCCQVCEAVR   EEVLADVRTIVQISYTPQDPPDLCGRL   YMASKNSSQETCTRARELAQQIGSHHISL   PAVKAVMGIPSLVTIKSPLFAAHGGSSR   ALQNVQARRMVLAYLFAQLSLWSRGV   LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIGREDJEAGMTY AELSVY   RKVAKMGPYSMFCKLLGMVRHICTPRQ   KVKRFSKYSMNRKKMTLTPAYHAEN   DNRFDLRPFLYNTSWPWQFRCIENQVLQ   AEPQSLDGVD   A 5741   1 5976   PGCAARLSRARAPGPGAAGAGRKRLAD   PASRRIRAPGGSRRAACAASPPPFOTPPRCP   LLLLLGAARAGALEIQRRFPSPTTPTNNFA   AAGTVYLAAVNRLYQLSGANLSLEAAR   AAGTV					}	Ì	1 = -
CIWKGYGKVCYKSATYPYIMLILFF TLPGAKEGILFYITPNFRKLSDSEVWLDA IFFSYGLGLGSLIALGSYNSFHNNYRDSI CINSCTSMFAGFVIFSIVGFMAHVTKRSIA AASGPGLAFLAYPGAVTQLISPLWALIFF LLMLGIDSQFCTVEGFITALVDEYPRLLRI ELFIAAVCISYLIGLSNITQGGIYVFKLFD ASGMSLLFLVFECVSISWFYGVNRFYDN MYGSRPCIWWKLCWSFFIPIIVAGVFIFS, MTPLTMGNYVFFKWGQOYGWLMALSSI IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDFGRTPELRGSAFRKTGANMPVRRG PONTFLGTIIRKFEGQNKKFIIANARVQNK YCNDGFGEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEJISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRITVNQISYTPQDPRDLCGRLI- YYMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTIGKSPLFAAHGGSSR ALQNVQARIRMVLAYLFAQLSLWSRGV LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAFA LEPHADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTILTPAYHAEN DNRFPLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGFGAAGAGKRLAD PASARRIARPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFOTPPRCPL LLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA			1	ŀ	İ		
TLPGAKEGILFYITPNFRKLSDSEVWLDAL IFFSYGLGLGSLIALGSYNSFHNNYRDSI CINSCTSMFAGFVIFSIVOFMAHVTKRSIA AASGPGLAFLAYPEAVTQLPISPLWAILFF LLMLGIDSQFCTVEGFTTALVDEYPRLLRI ELFIAAVCLISYLIGLSNITQGGTVYFKLFD ASGMSLLFLVFFECVSISWFYGVNRFYDN MVGSRPCIWWKLCWSFFTPILVAGVFIFS, MTPLTMGNYVFPKWGQGVGWLMALSSI IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRIPELRGSAPRKTGANMPVRRG PQNITFLGTIIRKFEGQNKKFILANARVQNK YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPPVSEE  728 2078 A 5737 1899 649 IQASRASPYRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRL YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAASPPPFOTPPRCPY LLLLLGAARAGALEIQRRFPSPTPTNNFR		1					
IFFSYGLGLGSLIALGSYNSFHNNVYRDSI   CINSCTSMFAGFVIFSIVGFMAHVTKRSIA     AASGPGLAFLAYPEAVTQLPISPLWALLFI     LLMLGIDSQFCTVEGFITALVDEYPRLLRI     ELFIAAVCLISYLIGLSNITQGGIVYFKLFD     ASGMSLLFLVFFECVSISWFYGVNRFYDN     MVGSRPCIWWKLCWSFFTPIVAGVFIFS     MTPLTMGNTVFFRWGQGVGWLMALSSI     IPGYMAYMFLTLKGSLKQRIQVMVQPSE     RPENGPEQPQAGSSTSKEAYI     PRRDPGRIPELRGSAPRKTGANMPVRRG     PRRDPGRIPELRGSAPRKTGANMPVRRG     PRRDPGRIPELRGSAPRKTGANMPVRRG     PRRDPGRIPELRGSAPRKTGANMPVRRG     PRRDPGRIPELRGSAPRKTGANMPVRRG     PRRDPGRIPELRGSAPRKTGANMPVRRG     PONTFLGTIRKFEGQNKKFILANARVQNK     YCNDGFCEMTGFSRPDVMQKPCTCD     VYNDGFCEMTGFSRPDVMQKPCTCD     LVKSPNVHYVYVLLLVLSGLLFYIPLHFI     AWFEKMTCYLQLLFNICLPDVSEE     10ASRASPYPRVKVDFALSCHEDLLAPIS     WKYHSPEEEISLGPACWLWDFLRRSQQA     LPLSGGVDSAATACLIYSMCCQVCEAVR     EEVLADVRTIVNOISYTPODPROLCGRL     YMASKNSQETAAHCQIGSHHISL     PAYKAVMGIFSLVTGKSPLFAAHGGSSR     ALQNVQARRMVLAYLFAQLSLWSRGVI     LVLGSANVDESLLGYLTKYDCSSADINI     ISKTDLRAFVOFCIQRFQLPALQSILLAFA     LEPLADGQVSQTDEEDMGMTYAELSVY     RKVAKMGPYSMFCKLLGMWRHICTFRG     KVRFFSKYSMNRIKMTTLTPAYHAEN     DNRFDLRPFLYNTSWPWQFRCIENQVLQ     AEQSLDGVD     T29   2079   A   5741   1   5976     PGCAARLSRARAPGFGAAGAGRKRLAD     PASRRLRAFGSRPRLAPCTRRAAQPAHA     PRASGRLRAPGSRPRLAPCTRRAAQPAHA     PRATTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOT			ł			İ	
CINSCTSMFAGFVIPSIVGFMAHVTKRSIA AASGPGLAFLA YPEAVTQLPISPLWAILFI LLMLGIDSQFCTVEGFITALVDEYPRLLRI ELFIAAVCIISYLIGLSNITQGGIVFKLFD ASGMSLLFLVFFECVSISWFYGVNRFYDD MYGSRPCIWWKLCWSFFTPIIVAGVFIFS, MTPLTMGNIVFPKWGQGVGWLMALSSI IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PQNTFLGTIIRKFEQQNKKFILANARVQNC VCNDGPCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYLLLVLSGLLFYPPLIHFI AWFEKMTCYLQLLFINICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSR ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMMFKKLLGMWRHICTFRQ KVKRFFSKYSMFKKLLGMWRHICTFRQ KVKRFFSKYSMFKKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHCTFRQ KVKRFFSKYSMFKKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMWRHICTFRQ KVKRFFSKYSMFCKLLGMRYRHSTPTPTNFA ALGRALARAGALEIQRRFPSPTPTNNFA AGGTVYLAAVNRLVQLSGANLSLEAEA	İ	1				1	
AASGPGLAFLAYPEAVTQLPISPLWAILFF LLMLGIDSQFCTVGEFITALVDEYPILLIST ELFLAAVCIISYLIGLSNITQGGIYVFKLFD ASGMSLLFLVFFECVSISWFYGVNRFYDN MVGSRPCIWWKLCWSFFFPIIVAGVFIFS MTPLTMGNYVFPKWGGGVGWLMALSSI IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PQNTFLGTIIRKFEGQNKKFILANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSFNVHYVYVLLLVLSGLLFYPPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASFYPRVVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL. YMASKNSSQETCTRAFELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPSWFKCKLLGMWRHICTTPQ KVKRFFSKYSMMRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASSPPPTPTNNFA AGGTVYLAAVNRLYQLSGANLSLEAEA		İ	1	•			
LLMLGIDSQFCTVEGFITALVDEYPRLLER ELFIAAVCIISYLIGLSNITQGGIYVFKLFD ASGMSLLFLVFFECVSISWFYGVNRFYDN MVGSRPCIWWKLCWSFFTPIIVAGVFIFS, MTPLTMGNYVFPKWGQGVGWLMALSS; IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PQNTFLGTIIRKFEQNKKFILANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYYYVLLLVLSGLLFYPLHFFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRL YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSR ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGFGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASSPPPTQTPPRCPY LLLLGAARAGALEIQRRFPSPTPTNNFA AGGTVYLAAVNRLVQLSGANLSLEAEA	1			1		1	
ELFIAAVCIISYLIGLSNITQGGIYVFKLFD ASGRSLLFLVFFECVSISWFYGVNRFYDN MYGSRPCIWKILCWSFFTPIIVAGYFIFS- MTPLTMGNYVFPKWGQGVGWLMALSSI IPGYMAYMFITLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PREDPGRTPELRGSAPRKTGANMPVRRG PQNTFLGTIIRKFEGQNKKFIIANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFKLCSFQVFLSFPLATIVIDVGLV- LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL- YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTIGKSPLFAAHGGSSRI ALQNVQARRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYHTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVRKFSKYSMNRHKMTILTPAYHAEN DNRFDLRFFLYNTSWPWQFRCIENQVLQ AEPQSLDGVU  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPFTTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA AAGTVYLAAVNRLYQLSGANLSLEAEA			1			1	
ASGMSLLFLVFFECVSISWFYGVNRFYDN MYGSRPCIWWKLCWSFFTPIIVAGVFIFS, MTPLTMGNYVFPKWGQGVGWLMALSS) IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PONTFLGTIIRKFEGQNKKFIIANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSPQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRLI, YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGHSLVTIGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVU LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVOFCIQRFQLPALQSILLAPA LEPPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRFFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAAASPPFQTTPRNFA AAGTVYLAAVNRLYQLSGANLSLEAEA AAGTVYLAAVNRLYQLSGANLSLEAEA	1		i		1		
MVGSRPCIWWKLCWSFFTPIIVAGVFIFSA MTPLTMGNTYPFK WGQCVGWLMALSSI IPGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PONTFLGTIIRKFEGQNKKFIIANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPODPRDLCGRIL YMASKNSSQETCTRARELAQQIGSHIISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAAASPPPFTTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEAA	1	1	1	1	İ		
MTPLTMGNYVFPK WGQGVGWLMALSSI   IPGYMAYMFLTLKGSLKQRIQVMVQPSE   RPENGPEQPQAGSSTSKEAYI     726   2076			1			ļ	
PGYMAYMFLTLKGSLKQRIQVMVQPSE RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PQNTFLGTIIRKFEGQNKKFIIANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLVLLSGLLFYPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  1QASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRIRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPFGTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.						1	MVGSRPCIWWKLCWSFFIPHVAGVFIPSAVQ
RPENGPEQPQAGSSTSKEAYI		1			İ	•	MIPLINGNIVER TO COLUMN CONTOUR
726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRG PQNTFLGTIIRKFEGQNKKFIIANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRILTYMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPFQTPPRCPV LLLLLGAARAGALEIQRRPPSPTPTINNFA AAGTVYLAAVNRLYQLSGANLSLEAEA			į	1			
PQNTFLGTIIRKFEGQNKKFIIANARVQNC YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL. YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTINIFA AAGTVYLAAVNRLYQLSGANLSLEAEA				<u> </u>		100	RPENGREUP AUSSISKEATI
727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNIPLIHFI AWFEKMTCYLQLLFNIPLIHFI AWFEKMTCYLQLLFNIPLIHFI AWFEKMTCYLQLLFNIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAQPAHA PRAAGGAPLSARAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	726	2076	A	5711	156	423	
727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLV LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL. YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	1	i	i		1		
LVKSPNVHYVYVLLLVLSGLLFYIPLIHFI AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASFYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRIL. YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA					<u> </u>		
AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRILTYMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVV LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAASPPPFQTPPRCPV LLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	727	2077	A	5716	3	274	
728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISI WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRILTYMASKNSQETCTTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADD PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGÁPLSARAAAASPPPFQTPPRCPY LLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA					1	į	
WKYHSPEEISLGPACWLWDFLRRSQQA LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRILT YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGÁPLSARAAAASPPPFQTPPRCPV LLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA		<u> </u>			<u> </u>		
LPLSGGVDSAATACLIYSMCCQVCEAVR EEVLADVRTIVNQISYTPQDPRDLCGRILT YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPY LLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	728	2078	A	5737	1899	649	
EEVLADVRTIVNQISYTPQDPRDLCGRILT YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIPSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPFFQTPPRCPY LLLLILGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	1	1			1	1	WKYHSPEEEISEGPACWLWDFLRRSQQAGFL
YMASKNSSQETCTRARELAQQIGSHHISL PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA: PRAAGGAPLSARAAAASPPFFQTPPRCPY LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	İ	1			1		LPLSGGVDSAATACLIYSMCCQVCEAVRSGN
PAVKAVMGIFSLVTGKSPLFAAHGGSSRI ALQNVQARRMYLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADI PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPY LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	1				1		
ALQNVQARIRMVLAYLFAQLSLWSRGVI LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVYI RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD PGCAARLSRARAPGPGAAGAGRKRLADI PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGÁPLSARAAAASPPPFQTPPRCPV LLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	1		1				
LLVLGSANVDESLLGYLTKYDCSSADINI ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	1			<b>\</b>	1	1	
ISKTDLRAFVQFCIQRFQLPALQSILLAPA LEPLADGQVSQTDEEDMGMTYAELSVY RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA		1	1		1		ALQNVQARIRMVLAYLFAQLSLWSRGVHGG
LEPLADGQVSQTDEEDMGMTYAELSVYRKVAKMGPYSMFCKLLGMWRHICTPRQKVKRFFSKYSMNRHKMTTLTPAYHAENDNRFDLRPFLYNTSWPWQFRCIENQVLQAEPQSLDGVDAEPQSLDGVDAEPQSLDGVDAEPQSLDGVDAEPQSLDGVDAEPQSLDGVDAEPQSRPRLAPCTRRAAQPAHAPRAAGGAPLSARAAASPPPFQTPPRCPVLLLLILGAARAGALEIQRRFPSPTPTNNFAAGGTVYLAAVNRLYQLSGANLSLEAEA	1	1 .	ı				
RKVAKMGPYSMFCKLLGMWRHICTPRQ KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA: PRAAGGAPLSARAAAASPPPFQTPPRCPY LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.						i	
KVKRFFSKYSMNRHKMTTLTPAYHAEN DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA: PRAAGGAPLSARAAAASPPPFQTPPRCPVLLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.	}	1	]	1	1		
DNRFDLRPFLYNTSWPWQFRCIENQVLQ AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.		1	1	1		I	
AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD: PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.		1	1				
729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLAD. PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.			1				
PASRRLRAPGSRPRLAPCTRRAAQPAHA PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	1					1	
PRAAGGAPLSARAAAASPPPFQTPPRCPV LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA	729	2079	A	5741	1	5976	
LLLLLGAARAGALEIQRRFPSPTPTNNFA AAGTVYLAAVNRLYQLSGANLSLEAEA.			}				PASRRLRAPGSRPRLAPCTRRAAQPAHARMA
AAGTVYLAAVNRLYQLSGANLSLEAEA	1		1				
	1				1	ì	LLLLLGAARAGALEIQRRFPSPTPTNNFALDG
	1				1		AAGTVYLAAVNRLYQLSGANLSLEAEAAVG
	1		ļ	1	1		PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL
		1			}		QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV
	Ì	1	1			1	AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS
				1	1	1	TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF
PRNRSLEDHRFENTPEIAIRSLDTRGDLA						1	PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT

		T 8 2 .	LODA	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID		Met	SEQ			D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F-Phenylalanine, G-Glycine, H-Histidine,
nucl-	peptide	j	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	1	USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino	Q=Glutamine, R=Arginine, S=Serine,
uence		l	914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
	1		] ~	amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
{		1	1	residue of	sequence	Y=Tyrosine, X=Unknown,sup codon,
1	1	1		peptide	ļ	/-possible nucleotide deletion, \-possible
1 .	1	1	Į	sequence	<u> </u>	nucleotide insertion
		+		<del> </del>		FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHP
i	ĺ		1		1	SDPPPGAQSYAYLALNSEARAGDKESQARSL
į.	Į.	1	1			LARICLPHGAGGDAKKLTESYIQLGLQCAGG
i	1		İ	}	}	AGRIGOLYSRLVSVFPARERLFAVFERPQGSPA
}			1	1	1	ARAAPAALCAFRFADVRAAIRAARTACFVEP
1 .	1		.	1		APDVVAVLDSVVOGTGPACERKLNIQLQPEQ
Ì		ł	1			I DCGA AHLOHPI SILOPLKATPVFRAPGLISV
Į.	1	1	1	j		AVASVNNYTAVFLGTVNGRLLKINLNESMQ
1	ĺ		1 .		1	VVSRRVVTVAYGEPVHHVMQFDPADSGYLY
				1	1	LMTSHQMARVKVAACNVHSTCGDCVGAAD
1	1	1	į	1		AYCGWCALETRCTLQQDCTNSSQQHFWTSA
1		1	1	}	1	SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS
1						LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ
1	1	1		1	1	IAYCNLLPRDQFPPFPPNQDHVTVEMSVRVN
1			1			GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA
		Į.	1			QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD
į.		- (			}	QWPCFWCSQQHSCVSNQSRCEASITY IST QD
1			ł	İ		CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG
-	- 1		l		ļ	AALECSFGLEEIFEAVWVNESVVRCDQVVLH
		ł	- [			TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM
			1	1		VYNCAMGSPDCSQCLGREDLGHLCMWSDGC
1		1	ł			RLRGPLOPMAGTCPAPEIRAIEPLSGPLDGGT
	1	ł	1			LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP
l l	1	1	İ			DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE
1	1	1	į.	1	ļ	GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI
1	1 .	ı	1			HGNDLHVGSELQVLVNDTDPCTELMRTDTSI
	1	- 1	- [	ł	1	ACTMPEGALPAPVPVCVRFERRGCVHGNLTF
•	1	i	l l	l l	1	WYMONPVITAISPRRSPVSGGRTITVAGERFH
		1		<b>\</b>	1	MVONVSMAVHHIGREPTLCKVLNSTLITCPSP
l	į		1	1		GALSNASAPVDFFINGRAYADEVAVAEELLD
		- [			1	PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH
	l	ļ	ł	1	1	HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY
	İ	- 1	l l			RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ
	1	- (	(	1	i	LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV
	1	1	1			LLLLSVVALFVFCTKSRRAERYWQKTLLQME
	1	İ	ı		l	EMESQIREEIRKGFAELQTDMTDLTKELNRSQ
	1	Ì	l	l		GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT
	ì	}	1		(	LNSQGSSQAQETHPLLGEWKIPESCRPNMEE
1	1	1	i	l	Į	GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD
1	1	1	1	1	1	RCSLASLLTIALHGKLEYYTSIMKELLVDLID
	1	1		l	l	ASAAKNPKLMLRRTESVVEKMLTNWMSICM
1		- }	1		Í	ASAAKNYKLIVILKKI ESV VERIVILLI I TITISIONI
1	1	- 1			1	YSCLRETYGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG
	1	1	j	1	1	KAKI ILINEE WELKENIEAKANILIN VOLQUCU
-	l	1	1	i		MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY
	1	1	ļ	1		SQWPRAEDVDLEWFASSTQSYILRDLDDTSV
·	1	1	1.	}	1	VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD
1		1		}		NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH
1	1	1		1	l	ROSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF
	- 1	i		1	1	KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD
	ł		1	1		PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK
l	1	1	1	1	·	TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN
j	1		ļ	ł		KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE
1	1	İ		- 1		MNAHLAEESRKYQNEFNTNVAMAEIYKYAK
1				- {	1	RYRPQIMAALEANPTARRTQLQHKFEQVVAL
1	ĺ	1	Ì	}	1	MEDNIYECYSEA
				<del>-  </del>	292	OPSPLEHSHLETLOLLRTAQLPEQVSWPWGQ
730	2080	)   A	5744	3	494	VANGKGNORNMGSPOPSLLAFERNLELQIMG
	}	- 1		1		LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT
- 1	ì		1 .		1	LKD
					202	FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC
73	208	1 A	5747	7   1	382	FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT
						THE ADDRESS OF THE PARTY OF THE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	•		·	peptide		/=possible nucleotide deletion, \=possible
	1			sequence		nucleotide insertion
-	<u> </u>					LSFSLRSSRVSGRHWKNFALVPLLREASARD
	l .				ĺ	RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS
	1				İ	GEK
732	2082	A	5753	198	3	AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
		_				PGKLQAQLPCPSPVRFTSARIPPASRPQTKS
733	2083	Α	5754	2	2223	AAGPPGLEAEGRAPESAGPGPGGDAAETPGL
ļ	!	Į.				PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL
1			1			SVANCLGAQTVQAPAEPAAGKAEQGETSGR
		1	1			EAPEAPAVGREDASAEDSCAEAGASGAADG
ł		ì	<b>†</b>	l	ļ	ATAPKTEEEEEEETAEVGRGAEAEAGDLEQ
	1			·		LNRTSTSTKSAKSGSEASASASKDALQAMILS
	,	i	İ			LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN
	1	ŀ	-			LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK
}		1	1	ł	ł	GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM
		l		ł	1	DFSSMELDEALRKFQAHIRVQGEAQKVERLIE
1		1				AFSQRYCMCNPEVVQQFHNPDTIFILAFAIILL
ļ		}		1		NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG
	•		1 .		_	ADIPRELVVGIYERIQQKELKSNEDHVTYVTK
İ	1	1			1	VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV
		1				NKLQKQAAHQREVFLFNDLLVILKLCPKKKS
Į.	ì	1	1			SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV
l		ļ				TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE
	1	1	1		1	SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA
					1	QGDPQSKQGSPTAKREAALRERPAESTVEVSI
		1				HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR
		1	1		1	QQTPPLPPPPPTPPGTLVQCQQIVKVIVLDKPC
	1	1	1	1	i	LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG
		1	,		1	SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH
L						GHRYSSGSRSLV SSVMGDLVGQGLEEQIVARDENSWLIDGGTP
734	2084	Α	5788	8	362	IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR
Ì		1	1	ļ		KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT
l	1	1		1		
						RIDSKATALSPKLPDAKDKEESVA MVFSAVLTAFHTGTSNTTFVVYENTYMNITL
735	2085	A	5827	1	1257	
	1					PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST
}	1	1	İ	1	1	AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG
	1 '	Ì				NLVVCLMVYQKAAMRSAINILLASLAFADM
1		1	1 -	1		LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF
		1	-			FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR
		Į				AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY
	1	1	1	[		SFMGILNTLRHNALRIHSYPEGICLSQASKLGL
			1	1		MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC
1						MOTOKLI NYLEGKTEAAUTI IIFITLYALIA
		1	1	l		WAPFTTYSLVATFSKHFYYQHNFFEISTWLL WLCYLKSALNPLIYYWRIKKFHDACLDMMP
}		1	1	I		WELLI LABALINELI I I WALAAFADALLUMME
L					1.00	KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	Α	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKQFSRS
		1	1	1		DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP
						LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR
		1		1	1	RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH
	1		1		1	GLARRLTTARRPPASSEQAQQELFNELKPAV
}						DGANFIVNHMRDQNNYNEEKDSWNRVART
1	1	1		1	1	VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP
		1	1	1		QPFPGDPYSYNVQDKRFI
<b>.</b>	1	- [	_ L			
738	2088	A	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG
738	2088	A	5881	1	1160	LLDSSKLCDYENRFNTSKGGELPDRPAGVGV
738	2088	A	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG LLDSSKLCDYENRFNTSKGGELPDRPAGVGV YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS

		X/-A	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
NO: of	NO: of	поа	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence			ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence		1	914		of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł	ļ	amino acid	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	İ		residue of	sequence	/=possible nucleotide deletion, \=possible
	1	1	1	peptide	j	nucleotide insertion
	İ			sequence	<u> </u>	WCSQGADCITPGLYAMVGAAACLGGVTRMT
		l				VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA
	Ļ	Ì	]	ļ	1	DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL
	1	1	1		} .	AMDVMKPRRNDPLLTVLTQDSMTVEDVETII
	i	1	1			AMDYMKPKKNDPLETVETQDSMTVEDVETH
		1	ļ			SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE
	ì	i	1			NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK
		į.	1	ł	i	LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC
	ì		1	ŀ	,	LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI
	1	1		1		LFN
739	2089	A	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP
137	2007	11	1007-	<u> </u>		DQALQELRKVARINGHKEAKNLTIEVLMSSV
	İ		1	1		KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV
	1		1	1	ì	VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA
	1		į.		1	VDFLGRATTALLLSFLGRRTIQAGSQAMAGL
	i	1	ŀ	i		AILANMLVPODLOTLRVVFAVLGKGCFGISL
		ì	1			TCLTIYKAELFPTPVRMTADGILHTVGRLGA
		1	1	Į	1	MMGPLILMSROALPLLPPLLYGVISIASSLVVL
		1				FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE
	1	1	1	1	,	AFTVESTSLLEIVALHGAL
			7000	<del> </del>	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE
740	2090	A	5900	2	420	NLIILDTAKKHGYEVVDTFTTMGRYKEFLQG
				i		KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM
}		1	- t	1	l.	GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV
1		1	ı	(	ĺ	CSEILLSRMCANKRTM
		J				RMPESTLLIICENGYILEAPLPTIKQEEDDHDV
741	2091	Α	5910	3	412	VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER
	ì	1	1	1	1	ORELKEKIREERRNKLAAEMGEDGEKEFQEE
i	1	1	- }		1	EEEKEEEEEEPLPEIFIPSTPSPILCGFYSEPG
ł	1	1	1	1	1	
						KFWV
742	2092	A	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV
	1	-	1	İ		TOMGNDKSIKCEONLGHDTMYWYKQDSKK
ļ		1	ì	1	1	FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL
!	]		1	Ì		NLHINSLELGDSAVYFCASSQDTALQSHCIPV
		1	1			HKPPGSARKLQGSVCTCTQGSSLHSLMASDG
	- 1	İ	i			VPVC
743	2093	T <sub>A</sub>	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER
' "		1		1		RALSVQQRGGPAWSGSLEWSRQSAGDRRRL
1			-	1	1	GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA
İ		1		1		ADRARRERFIMNEKWDTNSSENWHPIWNVN
1	1	ı	1	1	1	DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF
		- 1	1	1	1	LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI
1		1		1		LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM
1	1	1	}	1	1	CKISGLVOGISVAASVFTLVALAVDRFQCVVY
1		1		1		PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH
				1		VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ
			1	i	1	EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF
1		1				RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI
}		)		}	1	VALLFILSWLPLWTLMMLSDYADLSPNELQII
1	1	- 1	1	1		NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG
1		1				FOEAFOLOLCOKRAKPMEAYALKAKSHVLIN
1	1	1	1	1		TSNQLVQESTFQNPHGETLLYRKSAEKPQQE
	1	ł	1	l'		
1		1.				LVMEELKETTNSSEI
744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
1						LYDYQGGRLGVARGAWYMEAPDIRQGDM
745	2095	A	5970	413	856	GAPHTDWAWAPTPMSGLGSGRGRQGTLASS
143	2093	1.	-,,,	1	ì	PLSLPLLLAGVTGILATELFDQMARPAACMV
	ı	1	i			CGALMWIMLILVGLGFPFIMEALSHFLYVPFL
	-1	1	Ì			GVCVCGAIYTGLFLPETKGKTFQEISKELHRL
						GVCVCGAIYTGLFLPETKGKTFQEISKELHRL NFPRRAQGPTWRSLEVIQSTEL

ODO TO	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of		1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	ĺ		ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		ì	914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Į.	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	}	1		peptide	sequence	/=possible nucleotide deletion, \=possible
		1			]	nucleotide insertion
		<u> </u>	<del> </del>	sequence	10/10	AOTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
746	2096	A	5971	3	1343	RCARHGACQRSCLASQDPYCGWHSSRGCVDI
		İ				
	}	}				RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
		Ì	1			SGPGDSAYGVRDLPPASASRSVPIPLLLASV
		ł				AAAFALGASVSGLLVSCACRRAHRRRGKDIE
	İ		l		}	TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
		1			1	VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE
	1	l	1		i	LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
1	ļ	1	1		ļ	GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
	_	i			1	EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
	1	}	}			APALLGGPSPRPHECASPLRLDVPPEGRCASA
	[					PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
İ	1	1	1			LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG
	1					RALKRYDVEKPQLSLKPPLVGPSSRQAVPNG
!	1	1		1		GRFNF
747	2097	A	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI
]		1				LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
	1	1				RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN
	1					DRVQALSYAQHTRQLISCGGDGGIVVWNMD
1	ł	ł		ľ		VERQETPEWLDSDSCQKCDQPFFWNFKQMW
1	1					DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
İ		Ì	1			MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
]					Į	HNIVHVHFDATRGWLLTSGTDKVIKLWDMT
1	1	1	1		1	PVVS
748	2098	A	6001	12	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
/40	2090	1	0001	-	1	CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
1				1		TMLLMLKLCTEVRVANELNARRRSFTAADS
1	1		-	1		KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
		ŀ	1	İ		OCVLAFTGVAGYITYLSIDSALFVETLGFLAV
	1	1	}	<b>\</b>		LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
	ì	1				WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
			l	1	[	DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
210			6002	12	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK
749	2099	A	0002	12	447	RARLADDLNEKIAQRPGPMELVEKNILPVDSS
1	Į.		1	1		VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
		1	i		}	
1		[				DQPASQESQGSAASPSEPKVSESPSPVTTNTP
<u></u>	1	<del>  </del>	1	<del></del>	1400	AQFASVSPTVPEFLKTPPTAD
750	2100	Α	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
			1	ł	1	WRWELRLRNYVPEDEDLNKRRVPQAKPDAV
		[				QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD
		1		1		WDLKRDVAKKLEKLLKRTQRAIAELIRERLK
					J	GQEDSLDSAVDAATEHKTC
751	2101	A	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF
				1		SHPDKLKRMSKSVPAFLQDESDDRETDTASE
			1			SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
		1	1			VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
1				i		HVFVAQCKDLAAADVKKQRSDPYVKAYLLP
1		1	1			DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
			l l			QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
		1				TWDWDNKQNKQLRWYPLKRKTAPVALEAE
1	}			•		NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
		-	_ [ .		1	KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
	1		į		].	KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC
1		1		1		VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
1	1	1				EVDWMDSTSEEVALWEKMVNSPNTWIEATL
}		1	l	1	1	PLRMLLIAKISK
		1			+	
750	2102		6020	102	1283	I KERSPEEDSANFILLIAMA I COMMARENIA
752	2102	A	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SOVGGLGRFOMLHLVFILPSLMLLIPHILLENF
752	2102	A	6028	108	1283	SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF AAAIPGHRCWVHMLDNNTGSGNETGILSEDA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWYYDQSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK
753	2103	A	6043	1	1470	KNLKEKA  DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	A	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059		1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRPTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKILYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTITSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG
756	2106	A .	6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	A	6063	54	419	ITPLGLGAADMCAFPWLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QILTMLLRSLQQPSASWPRDCSSSCSW IGISCPATIFYPMFSHSLIGIGEEYQLPYYNMV PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC LRAVLKLMSECWAHNPASRLTALRIKKTLAK
759	2109	A	6072	3	650	MVESQDVKI PGRRFRPAALEERAMEKLREKVPFQNRGKGT LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV HSSVADMQNMPAAVHALLTQPSLSAAPFAQ RYLGTLPSTGSTTLPQCHAGNATVW
760	2110	A	6077	3	730	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG LRGCMLIELPLRGRLQLEACGMRRKSLLTRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ NWIELLSGETWNPLKLHYQLRNVRERLAKNL VEKGVLTTEKQNFLLFDMTTHPLTNNIKQR LIKKVQEAVLDKWVNDPHRMDRRLLALIYL AHASDVLENAFAPLLDEQYDLATKRVRQLLD LDPEVECLKANTNEVLWAVVAAFTK
761	2111	A	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLFKIMRVGTASS
762	2112	A	6079	2	1558	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSFPLPKQYLDVSSQTDISG SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI QLAKLDSEAWPGVLDSERDRILLINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTOSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII IQASTMIISRVPNISVHLLHEPPALTNEMYCLV

بله الأسلة المبط المبط المبط المساسد

				Dundlated	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
EQ ID	SEQ ID	Met	SEQ	Predicted	nucleotide	DesAgnartic Acid, Fe-Glutamic Acid,
O: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
ucl-	peptide		in		corresponding	I=Isolencine, K=Lysine, L=Leucine,
otide	seq-		USSN	location	to last amino	M=Methionine, N=Asparagine, P=Proline,
eq-	uence		09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ence	ļ ļ		914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
	l l		!!	amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ			1	residue of	sequence	/=possible nucleotide deletion, \=possible
	1			peptide	İ	/=possible nucleotide deletion, (-possible
	1		1	sequence		nucleotide insertion
	<del> </del>					VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK
	1		1			THVTLHGTELCDESYPALLTDIPVGDLHPGEQ
	i i				1	LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE
	1 !		1		}	KETVCKCHKDETVTIETVFPFDVAVKFVSTKF
	!		1		ł	EHLERVYADIPFLLMTDLLSASPWALTIVSSE
	į .		i			THE APSMITTUDOLESOVDNVILOTGESASECE
	ļ '	1	!		1	CLQCPSLGNIEGGVATGHYIISWKRTSAMENI
	1	1	1	1		PIITTVITLPHVIVENIPLHVNADLPSFGRVRES
	1	i	i	Ì	j	LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG
	1	(	l	1	l .	LPVKYHLQNKI DLVQDVEISVEI SDAI WI GO
		l	1			LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS
	1	1	į			LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM
		1			1	DDTSIAAA
	10111	<del>   </del>	6093	1	1422	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP
764	2114	Α	0073	*	1	DDDADADPTI.GNNHOESPGWRCCRPTLRERN
	1	ł	ł	1	1	ALMENNELMADVHFVVGPPGATRTVPAHKY
	i	1	1	1 .	1	VIAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA
	1	1	1	1	{	AFLILKYMYSDEIDLEADTVLATLYAAKKYI
		1	Ì	1		VPALAKACVNFLETSLEAKNACVLLSQSRLF
	1	1	1	1		VPALAKACVNFLETSBEAKTING VEGGECETOR
	l .	1	1	1	ł	EEPELTQRCWEVIDAQAEMALRSEGFCEIDR
		1	i	ì	1	QTLEHVTREALNTKEAVVFEAVLNWAEAEC
		1	1	1	ł	KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE
		Ì	1	1	1	FANGAAQSDILTLEETHSIFLWYTATNKPRLD
		1	1	<b>!</b>		FPLTKRKGLAPORCHRFQSSAYRSNQWRYKG
	1	1	ļ	1		RCDSIOFAVDRRVFIAGLGLYGSSSGKAEYSV
	Į.	1	-			KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF
	- {	1	1	,	į	EHPVQVEQDTFYTASAVLDGSELSYFGQEGM
	i i	1	1	1	ļ	TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE
		ì	ı	1		
	ì		1			LIFYA
765	2115	A	6099	1	1150	SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF
105	2113	1	4	j		RPVKAPGTFHMVHGKCMCKHNTAGSHCQH
		-	l l	Ĭ		CAPLYNDRPWEAADGKTGAPNECRTCKCNG
	- {	1		ì	l	HADTCHFDVNVWEASGNRSGGVCDDCQHN
	1	٠.	- 1	1	i	TEGOYCORCKPGFYRDLRRPFSAPDACKPCS
	1	-	1	į .		CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA
i	1	1	i	1		GRRCDRCMVGYWGFGDYGCRPCDCAGSCD
Ì	1	Į.	1.	·		PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP
1	1		Ì	1	ļ	WEWEDAQGFSALLHSGKCECKEQTLGNAKA
<b>S</b>	1	1				FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK
1	l l	1		- 1	Į.	FUMKISI VLAIMUSAHURUITI DI HITRIK
ĺ	1	l	1	1		KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL
Ī	1				i	NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK
1	- [	1		1		PSI GRKVMDILKRECK
			6103	-   2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR
766	2116	Α	0103	1 -	1	CVLTERGLOLFEAKGTGGRPKELSFARIKAVE
1	1	1		1		CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW
		1		1		NAQITLGLVKFKNQQAIQTVRARQSLGTGTL
	i	1		1		
1	1	1	1	<u> </u>		VS PROGRESS PLOSEDOTRI LAGMOSLA
767	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLI
1 '3'	211/	1 "		1		MARYYIIKYADQKALYTRDGQLLVGDPVAD
1	i		i	Į.	ļ	NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC
1	1	1	1	ļ		LACVETEEGPSLOLEDVNIEELYKGGEEATRI
ļ	1	1	,	ł	1	TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ
1		1	1	l	1	PVQLTKESEPSARTKFYFEQSW
1	1	- 1				FILQAVLQLSSQEARYKAFGTCVSHIGAILAF
768	2118	A	6109	3	292	TILQAVLQLOSQEAR I RAFOTO STRUMENT
1,00		1		1	1	YTPSVISSVMHRVARCAAPHVHILLANFYLLI
1	}	1	1	}	- 1	PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
		<del></del>	6110		711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS
769	2119	A	0110	1 *	1	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
1 .02		1	1	1	ł .	ETSDSEMEMEAEHYPNGVLGSMSTRIVNGA
"	1		li i		1	E I SDSEMEMENTALITIES A COMBAND AND AND AND AND AND AND AND AND AND
100	1	- 1	ł	ı		KHEDLQTDESSMDDRHPRRQLCGGNQAATI

	CEO TO	3/24	CEC	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	nou	in No.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq- uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	ucuce	1	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
dence			7.7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ĺ		{	peptide	•	/=possible nucleotide deletion, \=possible
	}		1	sequence		nucleotide insertion
	<del> </del>	├	<del>                                     </del>			RIILFGRELQALSEQLGREYGKNLAHTEMLQD
		ļ	}	ļ		AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
		}	1	Ì		NSAILESQNLPKQPPLMLALGQASECLRLMA
		1				RAGLGSCSFARVDDYLH
770	2120	A	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
'' <b>'</b>		}		-	1	VAPWALKYMNRRASQMLLMFLLAICLLAIIF
ł		1	1		}	VPQEMQMLREVLATLGLGASALANTLAFAH
	ļ	1				GNEVIPTIIRARAMGINATFANIAGALAPLMM
	1		1	ļ	}	ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
İ	1	1	ſ	· ·		PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	A	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
'''		1				RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC
-		1		1	ļ	LTKEDTGWYWCGIQRDFARDDMDFTELIVT
	i	1		į		DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
1		1				RKADRSRTSILIICILITGLGIISVISHLTKRRRS
1		1	1			QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	A	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
1		1				TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV
		1	ļ	1		WYLLRKHWIANNLFGLAFSLNGVELLHLNN
1			1	1		VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
ľ	1				}	FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
		1				VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
	1	1			1	GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
1		i	1	1	1	ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
						EASASKGLEKKEK
773	2123	A	6161	3	1088	COPMLVTRKNHPKLLLRRTESVAEKMLTNW
1				Í		FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG
1		-		1	ŀ	PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV
	'	1	Ì			NPENENAPEVPVKGLDCDTGTQAKEKLLDA
ļ		j				AYKGVPYSQRPKAADMDLEWRQGRMARIIL
				ł		QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
1						ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
					ł	SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
1		1	1	i		LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
1		1			1	VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
				1	1	NPQFVFDIHKNSITDACLSVV
L				1	106	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
774	2124	A	6163	860	125	
1		1			1	SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
		1			1	GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
	1	1		1	1	DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
1		1	-	· l	1	NHTMVYDGFGPADLRQACAELSLWDHGALA
1			1		1	NROLGGTRLSLGTGSSYGLQVPWMDSTPEEK
		ļ		1	ļ	OLWOALLEOPCEWVDGLLPLRTNLAPRT
L						ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
775	2125	A	6191	2	392	YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
1				1	ļ	DMK\KSPEIISRRMTFAL*CYSLTFVRFAHYVQ
1		i				VPWNWLMLGCHTAVDFDOLISSMPCISHGMT
	1	1	1	ĺ		ASASAL
				<b>_</b>	927	FRGYWGVREAFTDASWSGGLGPGKPGMKIT
776	2126	A	6217	1	827	ROKHAKKHLGFFRNNFGVREPYQILLDGTFC
		Ì	}	}	1	OAALRGRIQLREQLPRYLMGETQLCTTRCVL
1	1	1	1			KELETLGKDLYGAKLIAOKCOVRNCPHFKNA
	1	- 1		1	1	
	1	1		1	- {	VSGSECLLSMVEEGNPHHYFVATQDQNLSVK VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
1		- 1	}			VESG/RLSQCMRKKVSNISKRNRV**KTLNRG
			1			RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
-	1		1			KKRKRKRIRNRSNPKVLSEKQNAEGE
Ţ	1		l			LANGUAGE A LOSEN CHARGE

. ..... um som sam sam samb

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, B=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of peptide sequence	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL* FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/ SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC NSFRYRR
779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP YGQSQPSCFDRVKMGFVMGCAVGMAAGAL FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH QSQPMY
780	2130	A	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG TDYWLYSRGVCRTKSTSDNETSRKNEEVMT HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI S\ANAGRTPGQR\DSKKSYSGWSF/YFSGAFS FIIGRIIC*GVGLPWHIYIEKHQQLRAKSHSEF LKKSTFARLPPYRYRFRRSSSRSTEPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV
781	2131	Α .	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN
782	2132	A	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERKRESKILASFREMVHKKTKGKDDK
783	2133	A	6305	201	1032	WDDYPQGALRRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

NO: of mole- ceride seq- uence    NO: of mole- periodic contide seq- uence   1944   mucleotide     1944   mucleotide     1944   mucleotide     1945   mucleotide     1945   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1946   mucleotide     1947   mucleotide     1947   mucleotide     1948   mucleotide     1940   mucl	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
ende seg- uence 09/496 914 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
994% optopulation of the first summarian acid residue of peptide residue of peptide residue of peptide residue of peptide sequence per sequence seq							
uence    914   ng to first amino acid erisdue of peptide peptide peptide peptide peptide sequence   Thronian, V-Valine, W-Thytophan, Y-Tyrosiae, X-Uskawow, **Stop codon, /-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicion, V-possible uncleotide delicione (V-possible uncleotide delicion)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione)   V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide delicione (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-possible uncleotide (V-p			1	1			
mino acid of peptide residue of peptide residue of peptide sequenc		GOLIDO			•		O=Glutamine, R=Arginine, S=Serine,
residue of peptide   pep	20					of peptide	T=Threonine, V=Valine, W=Tryptophan,
			Ì	i			Y=Tyrosine, X=Unknown, *=Stop codon,
RSKLESLCRELQRINGSLKSEGVQFAREREE  KREVTSHFQVTINDIQLOMEQINITENSKIR QENMELABLIKKLEQYELREBIDKVFKIK DLQQQLVDAKLQQAGMIKARERERINGKO FILKBAVESQRMCEMKQQEHILKQQLAIY TEKFEFQNTI SKSEVFTITRQBMEMTIKKI KKLEKETIMYRSRWESSNKALLEMAEEKTV RDKELGQLQVLGRELEKLCZALQTGAQPTVR GQRWGSIRTSAVRIFS GQRWGSIRTSAVRIFS GGRWGSIRTSAVRIFS GGRWGSIRTSAVRIFS GGRWGSIRTSAVRIFS GGRWGSIRTSAVRIFS GGRWGSIRTSAVRIFS GGRWGSIRTSAVRIFS GDFLVHVDVGWQPHLGPTVGLACGRWT TLPOKUMORPMIKIPSEEVENSITSTISHDAI- KKCTQPAKPLAMRIVGSSVSPGFLVKWNNT REFERNSGTRVVSSCCOMSCAYSTICHGSVS ODEPLVHVDVGWQPPLGPTVGLRCGLLPHD TTPCQKLVVDDIDWA TREGRUKAVDDIDWA TREGRUKAVDDIDWA TREGRUKAVDDIDWA REFERNSGTRVSSCCOMSCAYSTICHGSVS ODEPLVHVDVGWQPFLGPTVGLRCGLSGVA GDFLVHVDVGWQPFLGPTVGLRCGLGSVA GDFLVHVDVGWQPFLGPTVGLRCGLGSVA REFERNSGTRVSSCCOMSCAYSTICHGSVS ODEPLVHVDVGWQPFLGPTVGLRCGLGSVA REFERNSGTRVSSCCOMSCAYSTICHGSVS ODEPLVHVDVGWQPFLGPTVGLRCGLGSVA REFERNSGTRVSSCORGATSTICHGSVA TREGRUKAVDDIDWA REFERNSGTRVSSCORGATSTICHGSVA REFERNSGTRVSSCORGATSTICHGSVA REFERNSCAVATA REFERNSGTRVSSCORGATSTICHGSVA REFERNSCAVATA REFERNSGTRVSSCORGATSTICHGSVA PRIAGESTORTALAN PRIAGESTORTALAN SGLGLAVWANTOVFGFSFLLITVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMVLFAGLTVALLASYS VHILLSMCGTATLGPTNTFMTSCHAGLTVALLASYS VHILLSMCGTATLGPTNTFMTSCHAGNTTMVLFAGLT TATTCYKKKMCGGMDFWLANTLFNTKGTATLANTV LAGGEDFVLTTGTTTTTTATTCYKKMGTGTGTATLATT TATTCYKKKMGCGGMDFWLANTLFNTKGTATLANTV LISCELMTANTVALGTCATATLATT TATTCYKKKMGCGMTSTCSCHAGNTATLATT TATTCYKKKMGCGMTSTCSCHAGNTATLATT TATTCYKKKMGCGMTSTCSCHAGNTATLATT TATTCYKKKMGCGMTSTCSCHAGNTATLATT TATTCYKKKMGCMTATLATT TATTCYKKKMGCMTATLATT TATTCYKKKMGCMTATLATT TATTCYKKKMGCMTATLATT TATTCYKKKMGCM		•	ì		peptide	-	/=possible nucleotide deletion, \=possible
RKREYTSHFQYTLNDIQLQMEQHARBINSKLR   QEMMELREHLKILEGYEYELREHIDKYKPKIK   DLQQU_VDAKLQQAQEMLKEAEERIQEEKD   FILKBAYESQRMEEMKKQETHIRLQQALIY   TEKFEFEQNTISKSSEYFTIFKQEMERMTKKI   KELKEKTIMYRSRWESSNKALLEMAREKTV   RDKELEGI_QVKIQRLEKL_CRALQT/GAQ*PVR   GQRWGSRTSANETS   GARGAGANETS   GARGAGA			į		sequence		
DiQQQLVDAKLQQAGEMLKGAERHOREMS	ļ		]	ļ	}	ļ	
FILIKEAVESQENCELMKQQETHLKQQLAU;   TEKPEFQATILISKSEVFITHKQEMEMMIKKI   KKLEKETIMYRSRWESSIKALLEMABEKTY   TEKPEFQATILISKSEVFITHKQEMEMCHKLK   KKLEKETIMYRSRWESSIKALLEMABEKTY   RDKELEGIQVKIQRIEKLCRALQITGQAPPVR   GQRWGSHRTSAVRIPS   SPQPDLIESVSPYSAGASSYPPGGAQPGYTTT   PSELVAVAPAGSAAGPAAGWQ+HACCRYNT   KLPWSWGMRPMKIPFSEEYRSISTISHDALP   KLPWSWGMRPMKIPFSEEYRSISTISHDALP   KLPWSWGMRPMKIPFSEEYRSISTISHDALP   KLPWSWGMRPMKIPFSEEYRSISTISHDALP   KLPWSWGMRPMKIPFSEEYRSISTISHDALP   KLPWSWGMRPMKIPFSEEYRSISTISHDALP   KLPWSWGMRPMKIPGSSVSPOPILVKWMWT   RREFRINGGTVVSSCOMSCM*SFLGHCSVS   QDLPLYHDVGWQPGMCGCTGEGPGNCKECISGYA   REHGQCADVDECSLAEKTVOKKNENCYNTP   GSYVCVCPDGFETRRCLCAAGREFGLEPKL   RHOPVACCDSSCVCTOEGPGNCKECISGYA   REHGQCADVDECSLAEKTVOKKNENCYNTP   GSYVCVCPDGFETRRCLCAAGREFGLEPKL   RHOPVACCDSSCVCTOEGPGNCKECISGYA   REHGQCADVDECSLAEKTVOKKNENCYNTP   GSYVCVCPDGFETRRCLCAAGREFGLEPKL   RHOPVACCDSCLAEKTVOKKNENCYNTP   GSYVCVCPDGFETRRCLCAAGREFGLEPKL   RHOPVACCDSCLAEKTVOKKNENCYNTP   GSYVCVCPDGFETRRCLCAAGREFALPKL   RHOPVACCDSCLAEKTVOKKNENCYNTD   GSYVCVCPDGFETRRCLCAAGREFALPKL   RHOPVACCDSCLAEKTVOKKNENCYNTD   GSYVCVCPDGFETRRCLCAAGREFALPKL   RHOPVACCDSCLAEKTVOKKNENCYNTD   GSYVCVCPDGFETRRCLCAAGREFALPKL   GSYSVALDGGTALLINGGUPTALTALLASYS   VILLLISMCIGTAYLGFTTNYFMVLPAH-LTCL   LUCKTON GROWN TO ALLING GRO		İ		ł		ł	
TEKPEEPQNTLSKSSEVFITTRQEMEKMTKKK   KELEKETIMYTSRWESSIKALLEMAEEKTV   RNKELEGLQVKIQRLEKLCRALQT/GAQ*PVR   GQRWGSHRTSAVRIPS							
	1		1	1		ļ	
RDKELEGLQVKIQRLEKLCRALQT/GAG*PVR	]	]	1	į.	ļ	ł	
R85			ļ		Ì	•	
1493   1493	ŀ	•			Ì	Ì	
PPSI.VAVAPAGSAGPAGWQ*HAGCGWT   KLPWSWGMRPKIFFSEPRISITBIDAL-  EKCTOPARPLSMRNTGSSVSPGPLVKWNWT   REREPNISOTRAYSCCMSWSPGJHCSWS   QDLPLVHYDVGWQPPLGPTVGLRPGLPLHD   TTPCQKLVYDDLDWA   TTPCQKLVVDDLDWA   TTPCQKLVVDDLDWA   TTPCQKLVVDDLDWA   TTPCQKLVVDDLGWA   REHGQCADVDECSLAKTCVRKIBICYNTP   GSYVCVCPDGETIRRCLAAGRG*SHRRK   PDTAALPRRPWMCRTYPLMYSEGCPVENVAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   RAMPSAVDSGGRIPAL   VAGTILIONIGAMSSYLLIKTELPAAIAEFLT   GDVSXWYLDGGTLLILICVGIVFFLALLERIG   FLGYTSSLSFFRMFFALVVIIKKWSIPCPLTL   NYVEKGFQISNVTDDCKPKLPHISKESAVALP   TMAFSICLHTSILPYYCELGSKRMONVTN   TAIALSFLIYFISALFGVILTYDGTTKAQRGE   VTCHRIKKVESELLKGG**IPSHIPSHDVVVMTV   KLCILFAVILTYPLHIPARKAVTMMFFSNFP   FSWRHPLITLALIMIVLLAIPYDRIVFGWGEDVVVMTV   KLCILFAVILTYPLHIPARKAVTMMFSNFP   FSWRHPLITLALIMIVLLAIPYDRIVFGWGED   RAMPSTALITYPSHDVVVMTV   KLCILFAVILTYPLHIPARKAVTMMFSNFP   FSWRHPLITLALIMIVLAIPHSHDVVVMTV   KLCILFAVILTYPLHIPARKAVTMMFSNFP   FSWRHPLITLALIMIVLAIPHSHDVVVMTV   KLCILFAVILTYPLHIPARKAVTMMFSNFP   FSWRHPLITLALIMIVLAIPHSHDVVVMTV   KLCILFAVILTYPLHIPARKAVTMMFSNFP   FSWRHPLITLALIMIVLAIPHSHDVVVMTV   KASHVPNOASLEOCHONLFCLADLTGIKWKK   YVWQGFTSAFIPTYTEEDFISSSCILKADV   KASHVPNOASLEOCHONLFCLADLTGIKWKK   YVWQGFTSAFIPTYTEEDFISSSCILKADV   KSSKLSCSFTFTHHGDSNVCTSVEINQHOPTV   LISEHTITAQOSNSPQVILCPFGLNGTLTGQ   AFKMSDSATKKLIGEWKGFYPISCLKEMSS   EKQEDMDWEDDSLAAVETVAGVRMTPAC   FVLVPQSDIPTSPVGSTHCSSSCILKADV   VRSSVSDGFNSDSTSHGGKPRKLANHVV   DRVWQECNMRAQNKRYSSGGICEEAT   AAKVASVDFVGATGRTNCSCILRIFKKSRN   AGQQQAQAPSLQQQQDLPKKTNSKNS   AGQQQAQAPSLQQQQQULFKTNCSCIRRIFKLKSRN   AGQQQAQAPSLQQQQQULFKTNCSCIRRIFKNCKSSFPQ   QRRPLTPFHHRVSVSDDVGMDADSASQRL   VNSSPDGVRSNRTNTDNARQOMKRYSPQMHGTE   MANSPQPPLSPHRCDVVDGVYKTPSTPQS   QHFYQMTTDDLVSRXMEDDANIAWKYYK   FYKKKDVETLPPQLESDKKNDDPVGTFQGESV   TYSTELMVQCKKPLKVSDELVQQYQKKQCSNCT   STSTELMVQCKKPLKVSDELVQQYQKKQCSNCT   STSTELMVCKKFLKFKNENGESSE   STSTELMCACKKFLKYSDENGENTNOSCIRTHOSCORD   AGROCK PRICKTRON	785	2135	A	6319	1493	889	
REKCTQPAKPLSMIRITGSSVSPGPLVKWNWT RREFRINGTRVYSSCOGMSCMYSFLGRICSVS QDLPLVHVDVGWQPPLGPTVGLRPGLIPLHD TTPCQKLVVDDLDWA  786 2136 A 6320 551 135 RWLPVAECDSSCVGCTGEGPGNCKECISGYA REHGQCADVDECSLAEKTCVRKNEKCYNTP GSVYCVCPDOFFETTRRCLCAAGRG*SHRRRK PDTAALPRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL  787 2137 A 6330 1693 227 DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGLGLAYVMANTGVFGFSFLLLTVALLASYS VHLLISMCQTAYLOPTNYFMVLPAH*LTCL PLIFFLQSL*NSL*AVTSYEDLGLFARGLPGKL VVAGTIIQNGAMSSYLLIKTELPAALAFELT GDYSRYWYLDGQTLLIICVGIVPFLALLPKIG FLGYTSSLSFFFMMFFALVVIIKLWSIPCPLTL NYVEKGPQISNNTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRWQNVTN TAIALSFLIYFISALFGYLITYPD/GTTKAQRGE VTCHRIKDKVESELLKG**1PPSHDVVVMTIV KLCLLFAVLUTVPLHFFAKATVMMFFSNFF FSWRHFLITLALMIIVLLATYVPDIRNVFGVV GASTSTCLETPFGLFYLKAAVTMMFSNFF FSWRHFLITLALMIIVLLATYVPDIRNVFGVV GASTSTCLETPFGLFYLKASPGLSWKKLGV GCPC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 1 6622 PRSLCFSLWABAVLADGGLRRRRRLRGTM KSHLSCSTFFFLHGDSNVCTSVEINQHQPVV LLSEEHITLAQGNSRFQVILCFGDLITGTTGQ AFKMSDSATKKLIGEWKYPYSPECDEKPIN KSHLLSCSTFFFLHGDSNVCTSVEINQHQPVV LLSEEHITLAQGNSRFQVILCFFGLIGTLTTQ AFKMSDSATKKLIGEWKYPYSECCLKEMSE EKQEDMDWEDDSLAAVEVLVAGCHVPAS TRDPAMSSVTLTPPTSPEEVQTVDGSVQKW VKFSSVSDGFNSDSTSHHGGKDPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEAT AAKVASWDFVBATQRTINCSCLRRKNIKSRN AGQQQAPSLGQQQQLRFHKTNIEKQEKSEK PQKRPLTFFHHRVSVSDDVGMDASASQRL VISAPDSSQNFRSNIKTNIDVAKITPQMHGTE MANSPQPPLSFHIPCDVVDEGVTKITSTPQS QHFYQMPTPDFVLPSKFRNDDVEGPGGESV TSVTELMVQCKKPLKVSDBLVQQYQKKNQCL	1.00		''				PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT
RREFRINSGTRVYSSCCGMSCMYSFLGHCSV/S QDLPL-WHYDWGWQPPLGPYGLPGHD-LPLHD TTPCQKLVVDDLDWA  8 6320 551 135 RWLPVAECDSSCVGTGEGPGMCKECISGYA REHGQCADVDECSLAEKT-VKRNENCYNTP GSYVCVCPDGFEETRRCLCAAGRG*SHRRR PDTAALPRPYMCRTYPLSGGCPVENVAL RMPSPAVDSGGERLPAL  787 2137 A 6330 1693 227 DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGLGLA-VYMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSL*NSL*NATYSPEDLGLFAFGLPGKL VVAGTIIQNIGAMSSYLLIIKTELPAALAEFLT GDYSRYWYLDGOTLLIICSPFKKRWGNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDK VESELLKG*VFSHDVMTNV KLCILFAVLLATYPLHFPARKAYTMMFSNFF FSWIRHFLITLALNIIUVLLATYVPDIRNVFGVV GASTSTCLIFIFGGFYLKLSREDFLSWKSLGV GCC/LLSFKTSILRNSLSVYIILPASRRSIYFKI  788 2138 A 6351 I 6622 PRSLCFSLW-WEAAVLAGGRRRRELLERGTM SASSYPNGASLEDCHCNLFCLADLTGIKWKK YVWQGTSAPILFPYTEEDPILSSFSRCIKADV LG7VWRQDGPRRREL*FWGGEDPVLLTLF TMTYQKKKMECGRMDFFMNAVLCTSKAVH NLLERCLMNRNFYRIGKWRYPYEEDBLISSFSRCIKADV LG7VWRDQRFRRREL*FWGGEDPVLLTLF TMTYQKKKMECGRMDFFMNAVLCTSKAVH NLLERCLMNRNFYRIGKWRYPYEEDBLISSFSRCIKADV LG8PMSNSVTLTPPTSPEEVQTVVDPGVVKW VKFSSVSDGFNSDTSTHFIGGRIPRKLANHV DRVWQCCMNNRAQNKRKYSASSGGLCEAT AAKVASWDFVEATQRTNCSCLRHKNILKSRN AGQGQAPSLGQQQLIFKHKTNEKQEESEK PQKRPLTFFHHRVSVSDDVGMDAJSASQRL VISAPDSSQVRFSNRXTNDVAKITPGMHGTE MANSGQPPLSFHEDDVDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLYSIKKFRIDDFVQFFGGESV TSVTELMVQCKKPLKVSDBLVQQFQIKNQCL	İ		ĺ		Í		
786 2136 A 6320 551 135 RWLPVAECDSSCVGCTGEGPGNCKECISGYA REHGQCADVDECSLAEKTCVKKNECYNTP GSVVCVCPDOFFETRRCLCAAGRG*SHRRK PDTAALPRPYMCRTYPLNYSEGCPVENVAL RMSPAVDSGERLPAL  787 2137 A 6330 1693 227 DYVLTAELHRQRSPGVSFGLSVTAMAIMG SGILGLAVMANTGVFGFSLLITVALLASYS VHLLISMCQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSI*NSL*AVTSYEDLGLFAGLPGKL VVAGTIIQNIGAMSSYLLIKTELPAALAFELT GDYSRYWYLDOQTILIIICVGIVFFLALLEKIG FLGYTSIS.SFFFMFFALLYKIK WSIPCPLIT. NYVEKGPGISNYTDDCKFKL HHPSKESAYALP TMAFSFLCHTSILPIYCELQSFSKKRMQNVTN TAIALSFLLYFISALFGYLIFTVD/GTTKAQRGE VTCHRIKDKVESELLKG**IP*SHDVVMTIV KLCILFAVLUTYPLHIFPARAAVTMMFFSNFF FSWRHFILTIALNIIVLLATYVPDIRTNYFGVV GASTSTCLIFPGLFYLKJSREDFLSWKKLGV GCCC/LLSFKTSILRNSLSVYIILPASRKSIYFKI TAIALSFLLYTISALFGYLTHYD/GTTKAQRGE VTCHRIKDKVESELLKG**IP*SHDVVMTIV KLCILFAVLUTYPLHIFPARAAVTMMFFSNFF FSWRHFILTIALATIVILATYPDFINNFGVV GASTSTCLIFPGLFYLKJSREDFLSWKSLGV GCCC/LLSFKTSILRNSLSVYIILPASRKSIYFKI TAIALSFLLYFISALFGYLTHYFISALFGYLTHYFISALFGYLTHYFISALFGYLTHYFISALFGYLTHYFITTATYTGKKKMECGRMDFPMAVLCTSKAYH NLLERCLIMNRNFYRIGKWKYPSEKDEKPIN KSEHLSCSFTFFLIGDSNVCTSVEHQIGPVV LLSEEHITLAQQNSFFQVLCPFGINGTLTGQ AFKMSDSATKKLIGEWKGYPSCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVMTYPAC FVLVPQSDIPTPSPVGSTFLOSSCLGVHQVPAS TRDPAMSSVTLTPTSPEEVQTVDPGSVQKW VKFSSVSDGFNSDTSHIGKTRIKLKSRN AGQGQAFSBATCHSTKNLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGQGGAFSBATSHIGKTRIKLKSRN AGAGGGAFFPPOPPLJSFRKFRINDSVASTPPOPPLSFRIFTCDVVDEGVTKTFSTPQS QHFYVQMFPOPPLJSFRKFRINDSVASTPPOPPLSFRIFTCDVVDEGVTKTFSTTPQS QHFYVQMFTPPPLJSFRKFRIDDSLSQSFPPOP YQEAVEPTVVVGTAVNLEEDEANIAWKYYK FPKKKDVEFLIPPCLJFSBKKFRIDDFVGFFGGESV TYPTEVTETUT AND ANTON A	Į.	,	ļ		-		
TTPCQKLVVDDLDWA			}	ļ	]	j	
786 2136 A 6320 551 135 RWLPVAECDSSCVGCTGEGFONCKECISGYA REHGQCADVDECSLAEKTCVRKNENCYNTP GSYVCVCPDGFETRRCLCAAGRG*SHRRKK PDTAALPRRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGELPJAL  787 2137 A 6330 1693 227 DYVLTAELHRORSPGVSFGLSVFNLMMAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYLIG*TNAFTAH*LTCL PLIEFLQSL*NSL*AVTSYEDLGLAFGLPGKL VAGGTIUNGINGAMSYLLIKTELPAALAFELT GDYSRYWYLDGOTLLIUCVGIVPFLALLPKIG FLGYTSSLSFFFMMFFALVVIKKWSPCPLTI. NYVEKGFQISNVTDDCFYKLFHFSKESAYALP TMAFSFLCHTSILPTYCELQSPSKEMQNVTM KLCILFAVILATVPDIRTVACKGRGE VTCHRIKDKVESELLKG***IP*SHDVVMTNY KLCILFAVILATVPDIRTVACKGRGE VTCHRIKDKVESELLKG***IP*SHDVVMTNY KLCILFAVILATVPDIRTVACKGRGE VTCHRIKDKVESELLKG***IP*SHDVVMTNY KLCILFAVILATVPDIRTVACKGRGE VTCHRIKDKVESELLKG***IP*SHDVVMTNY KLCILFAVILATVPDIRTVFGVV GASTSTCLIFIFFGLFYLKLSREDDISWKKLGV GGCFC/LLSFKTSILRNSLSVYIILPASRRSIYFKI SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGFTSAPILPFVTEEDPILSSFSRCLKADV LG/WRRDQRERRELI*FUNGGEDPVLLTLF TMTYGKKMECGMDFFMNAVLCTSKAVH NLLSRCLMNRNFVRIGKWFVKPYERDEKPIN KSSHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLJGEWKGVFYSECCKEMSE EKQEDMWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDTSHHGGKUPRLANHVY DRVWQECNMRAQNKRXYSASGGLCEAT AAKVASWDFVEATGRTNCSCLRHKNLKSNN AGQQGAPSLGQQQQLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMADASVASQRL VUSAPDSQVFFSINKTNDVAKITSPMGTE MANSPQPPPLSPHPCDVDGEOVTKTFSTPQS QHFYQMPTPDLVPSKFMEDRISDSSPPQ YQEAVEPTVYVGTAVVALEEDBANIAWKYYK FPKKKDVEFLPPQLSDKKKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQVQRENKY			ł	1	Ì	Į	1 *
REHGQCADVBECSLAEKTCVRKNENCYNTP GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRR PDTAALPRRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL  787 2137 A 6330 1693 227 DYVLTAELHRQRSSPOSFGLSVFNLMNAIMG SGILGLAYVMANTOVFGFSPLLLTVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTC: PLIEFLQSL*NSL*AVTYSYEDLGLFAFGLPGKL VVAGTIIIQNIGAMSSYLLIIKTELPAALAEFLT GDYSRYWYLDGQTLLIVGGIVFPLALIPKIG FLGYTSSLSFFFMMFFALVVIKKWSIPCPLTL NYVEKGFQISNVTDDCKFLFHSSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYDIGSTKAYMOVTN KLCLLFAVLLIVVPLHHFPARRAAVTMMFFSNFF FSWRIFHEITLAALMIVLATYVPDIRNYFGVV GASTSTCLIFFPGLFYLKLSREDFLSWKKLGV GCC/ALSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 1 6622 PRSLCTSILWAEAVALADGGLRRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADTGIKWKK YVWQGFTSAPLIFPYTEDPILSSFSRCLKADV LGVWRRDQRFERFEL*IFWGGEDPVLLTLF TMYQKKKMECGRMDFFMNAVLCFSKAVH NLLBRCLMNRNFVRIGKWFVKFYFKDEKEND KSEFLSCSFTFFLHGDSNVCTSVENOHQPVY LLSEHITLAQQSNSFPQVILCFFGLNGTLTGG AFKMSDSATKKLJGEWKOFYFISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPTSPEEVQTVDPQSVQKW VKFSVSSDGFNSDSTSHHGGKUFRLLANHVV DRVWQECNMRRAQNKRKYSASSGGLCEAT AAKVASWDFVEATGNSCLRHKNLKSRN AGQQGQAPSLGQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRNSVSDDVGMDADSNASQRL VISAPIDSQVWFSINITNDVAKITPOMHGTE MANSPOPPPLSPHIPCDVVDGEVGKGESSFPQ VQEAVEPTVVVGTAVNLEEDEANLAWKYYK FPKKKDVEFLPPQLFSDKKKDDVGFFGGESSFPQ VQEAVEPTVVVGTAVNLEEDEANLAWKYYK FPKKKDVEFLPPQLFSDKKKDDVGFFGGESSFPQ VQEAVEPTVVVGTAVNLEEDEANLAWKYYK FPKKKDVEFLEPPQLFSDKKKDDVGFFGGESSFPQ VQEAVEPTVVVGTAVNLEEDEANLAWKYYK FPKKKDVEFLEPPQLFSDKKKDDVGFFGGESSFP			<del>  _   _     _</del>	7222	7.51	126	
GSYVCVCPDGFEET/RRCLCAAGRG*SHRRK PDTAALPRPWMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL  787 2137 A 6330 1693 227 DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGLGLAYVMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYL.GP*INYFMVLPAH-LTCL PLIEFLQSL*NSL\*AVTSYEDLGLFAFGLPGKL VVAGTHIQNIGAMSSYLLIJKTELPAALAEFLT GDYSRYWYLDOQTILLIJCVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIIKK WSPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESA, YALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLITYDJGTTKAQRGE VTCHRIKDK VESELLK.GF**IP\$*SHDVVVMTV KLCILFAVLLIVTPLHEPAAKAVTMMFFSNFP FSWIRHFLITLALNIIJVLLAIYVPDIRNVFGVV GASTSTCLIFFPGLFYLKLSREDFLSWKKLGV GCFC/LLSKKTSILRNSLSVYILPASKKSIYFKI GCFC/LLSKKTSILRNSLSVYILPASKKSIYFKI GASTSYVDGGFGRAFALVLADGGGRRRRRLLRGTM SASFVPNGASLEDGLCOLFCLALDIGKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV L.G/VWRRDQRPBRREU**IFWGGEDPVLTLIF TMTYGKKKMEGGBONVCTSVEINOHQPVY LLSEEHITLAQQSNSPFQVILCPFOLNGTLTGQ AFKMSDSATKKLIGEWKGYPISCCLKEMSE EXCEDDMWEDDSLAAVEVLVAGYMTYPAC FVLVPQSDIPTPSPVGSTHCSSCLGVHQVPAS TRDPAMSSVTLTPTSPEVQTVDPOSVOKW VKTSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMRAQNKRKYASSGGLCEEAT AAKVASVDFVFATGTRTCSCLRHKNILKSRN AGQQGQAPSLGQQQQLPKHKTNEKQEKSEK PQKRPLTPFHRVSVSDDVGMDADSASQRL VUSAPUDSQVRTPSRIRTINDVAKITPQMHGTE MANSPQPPFLSNHTCDVVDBGVTKTTSTTPQS QHFFYQMTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEFTYVYGTAVNLEEDBANTAWKYYK FFKKKDVEFFLPPQLPSDKFKNDDVCDFFGQESV TSVTIELMVQCKKFRLDDFVGPFGQESV TSVTIELMVQCKKFRLDPVQDFGVGENQCT	786	2136	A	6320	551	135	
PDTAALPRPYMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL  787 2137 A 6330 1693 227 DYVLTAELHRQRSPOVSFGLSVFNLMNAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHLLISMCQTAYLPAHLTCL PLIEFLQSL*NSL\*AVTSYEDLGLFAFGLPGKL VVAGTIIIQIMAMSYLLLIIKUFGLPTAHLTCLI GDYSRYWYLDOGTLLIIICVGIVPFLALLPKIG FLGYTSSLSFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSLLPYJCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLITYD/GTTKAQRGE VTCHRIKOKVESELKG***IP*SGIDVVVMTIV KLCILFAVLL\TVPLHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAIYPDIRNVFGOV GASTSTCLIFIFPGLFYLKLSREDFLSWKLCGV GCFCALLSFKTSLRNSLSVVIILPASRKSIYFKI 788 2138 A 6351 1 6622 PRSLCFSLWAEAAVLADGGLRRRRRLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/WRRDQRFBREL*FIFWGGEDPVLLTIF TMTYQKKKMECGRMDFPMAVLCFSKAVH NLLBRCLIMNRNFVRIGKWFKYPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQEVY LLSEHITLAQQSNSFPQVLCFFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCIKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVPQSDIPTPSFVGSTHCSSCLGVHQVPAS TRDPAMSSVTLTPFTSPEEVGTVDPGSVQKW VKFSSVSDGFNSDSTSHHGGKUPKLANHVV DRVWQECNMRAQNKRYSASSGGLCEEAT AAKVASWDFVBATGRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQUFTPSHVKTNEKQEKSEK PQKRPLTPFHREVATORSCLRHKNLKSRN AGQQGQAPSLGQQQUFKKTNEKQEKSEK PQKRPLTPFHREVATORSCLRHKNLKSRN AGQQGQAPSLGQQQUFKKTNEKQEKSEK PQKRPLTPFHREVATORSCLRHKNLKSRN AGQVGAPSLGQQQUFKTNCSCLRHKNLKSRN AGQVGAPSLGQQQUFKTNCSCLRHKNLKSRN AGQVGAPSLGQQQUFKTNCTSTTQS QHFYQMFTDPDI-VPSKPMEDRIDSS.QSFPPQ YQBAVEFTVYVGTAVNLEEDBANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGFGQESV TSVTELMVQCKKFLKVSDELVQQVQKRNQCL	1				•		
RMPSPAVDSGERLPAL  2137 A 6330 1693 227 DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGLGLAYVMANTGVFGFSFLLLTVALLASYS VHILLSMC(QTAYL,GP*INVYMVLPAH-LTCL PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL VVAGTII(QNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDQQTLLIILCVGIVFFLALLPKIG FLGYTSSLSFFMMFFALVVIIKK WSIPCPLTL NYVEKGFQISNVTDDCKFKLFHFSKESAYALP TMAFSFLGHTSLIPLEQSPSKKRMQNYTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKKN VESELLKG**IP-SIPSIDVVYMTV KLCILFAVILAVTPLHEPARKAVTMMFFSNFF FSWIRHFLITLANIIIVLLAIYVPDIRNVFGGV GASTSTCLFFFGGFYLKLSREDFLSWKLLGG GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 1 6622 PRSICFSLWAEAAVLADGGLRRRRILRGTM SASFVPNGASLEDGNIFCTSUCHOLFICLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/WRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMAVLCFSKAVH NILBRCLIMMRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLIGSDVCTSVERNHQPVY LLSEEHTILAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMYPAC FVLVPQSDIPTPSFVGSTHCSSCLGVHQVPAS TRDPAAMSSVTLTPPTSPEEVQTVDPGSVGKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMRAQNKKRYASSGGLCEEAT AAKVASWDFVEATORTINCSCLIFIKNILKSRN AGQQGQAPSLGQQQLPKHKTNEKQEKSEK PQKRPLTPFHRVSVSDDVGMDNADSASQRI- VUSAPUDSQVARFSNRIVINDVAKITPQMHGTE MANSPQPPFLSNHPCDVVDBGVTKTRSTTPQS QHFYQMFTDPDLVPSKPMEDRIDISS.QSFPPQ YQBAVEFTVYVGTAVNLEEDBANTAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFQGESV TSVTELMVQCKKFLKVSDELVQQVQIKNQCI			ļ	ļ ·	}		
787 2137 A 6330 1693 227 DYVLTAELHRQRSPGVSFGLSVFNIMNAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHILLSMCQTAYLGP*TNYFMVLPAH-LTCL PLIEFLQSL*NSL*N*AVTSYFDLGLFAFGLPGKL VVAGTIIQNIGAMSSYLLIKTELPAA1AFFLT GDYSRYWYLDGQTLLIIICVGIVFFLALLPKIG FLGYTSSLSFFFMMFFALVVIKKWSIPCPLTL NYVEKGFQISNYOTDDCKPKL*IPFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG**IP*SHDVVVMTV KLCILFAVLLVIVPLHFFARKAVTMMFFSNFP FSWIRHFLITLAUITVLADYYDFIRNVFGVV GASTSTCLIFIFFGLFYLKLSREDFLSWKLGV GCCCLLSFKTSILRNSLSVYIILPASRKSIYFKI PRSLCYSLWABAVLADGGLRRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPLFPVTEEDPILSSFSRCLKADV LGVWRRDQRPEREL**IFWGGEDPIVLLTLF TMTYQKKKMECGRMDFPMAVLCFSKAVH NLLERCLMRNFVRIGKWFVKYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSFFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVQSDIPTSFVGSTHCSSCLGVHQVPAS TRDPAMSSVTLTPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEAT AAKVASWDFVEATQRTINCSCLRHCNLKSRN AGQQGQAPSLGQQQQLFKHKTNEKQEKSEK PQKRPLTPFHHRQVSUDDGVMCADSASQRL VISAPDSQIVRFSNIRYINDVAKYTPQMHGTE MANSPQPPLSPHHCDVVDGGVTKTNSTRQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWSYK FPKKKDVEFLPPQLPSDKFKDDPVGPGESV			1	1	İ		
SGI.GLAYYMANTGVFGFSFLLITVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSI*NSL*AYTSYEDLGLFAFGLPGKL VVAGTIIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG**IP*SHDVVVMTIV KLCILFAVLLIVYPLHFPARKAYTMMFFSNFP FSWIRHFLITLALNIIVLLAIYVPDIRNVFGVV GASTSTCLFFFGLFYLKLSREDFLSWKLGV GCFC/LSFKYLSIRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 I 6622 PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCHLFCLADLTGIKWKK YVWQGPTSAPLFPVTEEDPILSSFSRCLKADV LG/WRRDQRPERREL*FFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQOSNSPFQVLLCPFOLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVPQSDIPTPSYGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQGCNMRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGAPSLGQQQQLKHKTNEKQEKSEK PQKRPLTFPHHRVSVSDDVGMDADSASQRL VISAPDSQIVRFSNIRITINDVAKTIPQMHGTE MANSPQPPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKFMEDRIDSLSQSFPPQ VQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGGESV TSVTELMVQCKKPLKVSDBLVQQYQIKNQCL	787	2137	A	6330	1693	227	
PLIEFLQSL*NSL!*AVTSYEDLGILFAFGLPGKL VVAGTIIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDOQTLLIIICVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIIKK WSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTYDJGTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMTV KLCLLFAVLLVTVPLHFPARKAVTMMFFSNFP FSWIRHFLITLAINIIVLLAIYVPDIRNVFGVV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWCSTCLAADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKLSEADFLSWKLGV GASTSTCLIFFGLFYLKNGLSWKLGV GASTSTCLIFFGLFYLKNGLSWKLGV GASTSTCLIFFGLFYLKNGLSWKLGV GASTSTCLIFFGLFYLKNGLSWKLGV GASTSTCLIFFGLFYLKNGLSFKADPV LGFVUPGSPFGLFYLTTGAT MANGSONSPFQUILCFGLNCTILTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVPQSDIPTPSPVGSTHCSSCLGWHQVPAS TRDPAMSSVTLTPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKLPKKLANHVV DRVVQECNMRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLQQQQLPKHKTNEKQEKSEK PQKRPLTPFHIRVSVSDDVGMDADSASQRL VISAPDSQVVRFSNIRVNDVAKITPQMHGTE MANSPQPPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMTTPDLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVQFFGQESV TSVTELMVQCKKPLKVSDELVQQYQKNQCL	] '''	1 213	1	5550	1000		
VVAGTÜIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDGQTILIIICVGIVFPLALLPKIG FLGYTSSLSFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFALALFGYLTFYDJGTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMTIV KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP FSWRIFILITALNIIVILIATYPDIRNVFGVV GASTSTCLFIFPGLFYLKLSREDFLSWKKLGV GCCC/LLSFKTSILRNSLSVYIILPASRKSIYFKI 788 2138 A 6351 ! 6622 PRSLCPSUMAAAVLADGGRRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGFTSAPILFPVTEEDPILSSFSRCLKADV LG/WRRDQFRRERL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMRNRFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDLAAVEUV VAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGKFJPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGAPSTORTNCSCLRHKNLKSRN AGQQGAPSTORTNCSCLRHKNLKSRN AGQQGAPSTPHFHRVSVSDDVGMDVADSASQRL VISAPDSQVRFSNRITNDVAKITPQMHGTE MANSPQPPLSPHRCDVVDEGVTKTPSTPQS QHFYQMPTDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWXYYK FPKKKDVEFLPPQLPSDKFKDDPVQFFQGESV TSVTELMVQCKKPLKVDSDLVQQYQIKNQCL					1		VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL
GDYSRYWYLDGQTLLIIICVGIVFPLALLPRIG FLGYTSSLSFFFMMFFALVVIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYDGTTKAQRGE VTCHRIKDKSESLLKG**-IP*SHEDVVMTIVV KLCILFAVLL\TVPLHIFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSI\YFKI  788 2138 A 6351 I 6622 PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPYTEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NILBRCLMRNFVVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSSEHITLAQQSNSFQVULCPFGLNGTLTGQ AFKMSDSATKKLIGEWKGYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHIGGKPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPPHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQVFSNIR\TNDVAK\TYCMHGTE MANSPQPPLSPHPCDVDEGVTKTFSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVETTVYGTAVILEEDEANIAWKYYK FPKKKDVEFLPPQLPSKFKDDPVGFGQESV STSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1					· .	
FLGYTS\$LSFFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYDGTTKAQRGE VTCHRIKDKVESELLKG***IP*SEDVVVMTVV KLCLLFAVLNTVPLHFFPARKAVTMMFFSNFP FSWIRHPLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFFGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788  2138 A 6351  6622  PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/WRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSFFQVULCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTILTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKUPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAFSGQQQQLIPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDADSASQRL VISAPUSQNVFFSNIRINDVAKKIPQMHGTE MANSPQPPPLSPHPCDVDEGVTKTFSTFQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVILEEDEANIAWKYYK FPKKKDVEFLPPQLPSKFNDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	Ì	1		1		
NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLYPISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMIVV KLCILFAVLL\TVPLHFFARKAVTMMFFSNFP FSWIRHFLITLALNIIVLLAIYVPDIRNVFGVV GASTSTCLFFIPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 I 6622 PRSLCFSLWAAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRFERREL*IFWGGEDPVLLTLF TMTYQKKKMEGGRMGDFMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLIGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMRRAQNKRKYSASSGGLCEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGAPSLGQQQLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDADSNASQRL VISAPUSQVVRFSNIRITNDVAKITPQMHGTE MANSPQPPPLSPHIPCDVVDEGVTKTTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANTAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGGES TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	1					
TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMTV KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVILAITYVPDIRNVFGVV GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV GCFC/LJSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 1 6622 PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDP/VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKYPEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDADSASQRL VUSAPDSQVVRFSNIRYTNDVAK\TPQMHGTE MANSPQPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK PPKKKDVFLPPQLPSDKFKDDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	ĺ		1	1 .	1	1	
TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMTIV KLCILFAVLLUTYPLHFPARKAVTMMFFSNFP FSWRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI PSLCFSLWAEAAVLADGGLRRRRILLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPYTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSFFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKLPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEAAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLLPKHKTNEKQEKSEK PQKRPLTFFHHRVSVSDDVGMDVADSVASQRL VISAP\DSQ\VRFSNR\TNDVAK\TPQMHGTE MANSSQPPPLSFHPCDVVDEGVTKTPSTFQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLYSDKFKDDPVOFFQGESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL		1			1		
VTCHRIKDK VESELLK G***IP*SHDVVVMT\V KLCILFAVLL\TYPLIHFPARKAVTMMFFSNFP FSWRHFLITLALNIIVLLAITYPDIRNYFGVV GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 I 6622 PRSLCFSLWAEAAVLADGGLRRRRRLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGFTSAPILFPYTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGK WFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQOSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLJKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADSVASQRL VISAPDSQVRFSNIRITNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQKAVEPTVYVGTAVNLEEDEANIAWKYYK FFKKKDVEFLPPQLPSDKFKDDPVGFFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	Į.						
KLCILFAVLLYTVPLIHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAITYVPDIRNVFGVV GASTSTCLIFFPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 I 6622 PRSLCFSLWAEAAVLADGGLRRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPYTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKLPRKLANHVV DRVWQECNMRAQNKRKYSASSGGLCEAT AAKVASWDFVBATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQULPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDVANSVASQRL VISAPDSQVVRFSNIRVINDVAKITPQMHGTE MANSPQPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	į.	1	1		1	
FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 I 6622 PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPYTEEDPILSSFSRCLKADV LG/VWRQDRPERRELL*IFWGGEDPIVLTLF TMTYQKKKMECGRMDFFMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAPDSQVVRFSNIRINDVAKKTPQMHGTE MANSPQPPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL			}	1			•
GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI  788 2138 A 6351 I 6622 PRSLCPSLWAEAAVLADGGLRRRRRLIRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFFMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADSNASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL			1				
788 2138 A 6351 I 6622 PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDPVLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSFFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQULPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	ł	1		1	1	
SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERREL*IFWGGEDPVULTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSFVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPKKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHIFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	i			]			
YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSVSDGFNSDSTSHHGGKIPKLANHVV DRVWQECNMNRAQNKRKYSASSGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQVRFSNIR\TNDVAK\TPQHHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FFKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	788	2138	A	6351	1	6622	
LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL			1	_			1
TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQLLPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	1		1			YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV
NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQULPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDNADSVASQRL VIISAPUSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL						1	
KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQULPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDNADSVASQRL VIISAPVDSQVVRFSNIRVTNDVAKYTPQMHGTE MANSPQPPPLSPYHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL				1			1 7
LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDNADSVASQRL VIISAPDSQIVRFSNIRVTNDVAKVTPQMHGTE MANSPQPPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL		j '	1	}		}	•
AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIRYTNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL			1	1	1		
EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTINCSCLRHKNLKSRN AGQQGQAPSLGQQQULPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TINDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL				1	ļ		**
FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTINCSCLRHKNILKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL						1	
TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADSVASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL			1	}			
DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMDIADSVASQRL VIISAPVDSQVVRFSNIR\TNDDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	1	1	1			
AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQULPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL				}		-	
AGQQGQAPSLGQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TINDVAK\TPQMHGTE MANSPQPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1.	1	1	1	1		
PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1		1	1	ł		
VIISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1	1	1	[	1		
MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL	1				1		
QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL		1	1	1 .		ļ	
YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL					J	}	
FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL		1	1		ļ		1 7 7
TSVTELMVQCKKPLKVSDELVQQYQIKNQCL					1		
					1	1	
	1		1		Ī		

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine O=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uchoc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			7,14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Sequence	/=possible nucleotide deletion, \=possible
					l .	nucleotide insertion
				sequence		RONSEREAGKKHKVEDGTSSVTVLSHEEDA
		1			ļ	KONSEREAGENER VEDGISSVI VESIDEDA
		İ			l	MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA
					ł	VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
			[ .			ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
!		ļ			ļ	EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
						GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
1	1				ļ	KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
İ						CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
į.	l	i	1			DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
	!	į.				LPSPSTPRFPTPRTPRTPRTPRGAGGPASAQGS
1	ł	l	1			VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
ļ	ļ	1				EAHSLYVNLILSESVMNLFKDCNSDSCCICVC
-	1	1			1	NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
İ	1	l	ļ '			MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
1	Ì	1	1		1	KRFEALRATSAEHVNGGLKESEKLSDDLILLL
Ì	ĺ	1	1			QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
			-		1	
İ		1		İ		EERDCCNDCYLALEHGRQFMDNMSGGKVDE
1	}	1	ł	:	1	ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
	1		ı	ļ		LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH
1	1			ļ		KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV
			į	·	1	LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
1			1			NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
1		1	1			LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
1	i .	1	1	ı		GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
1		İ	1			SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
ł		ł		!		ATLASAASSTMTVTSGVAISTSVATANSTLTT
1		ì	}	1		ASTSSSSSNLNSGVSSNKLPSFPPFGSMNSNA
i i		ŀ				AGSMSTQANTVQSGQLGGQQTSALQTAGISG
1					1	ESSSLPTOPHPDVSESTMDRDKVGIPTDGDSH
1	1		1	1		AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
		1	1			GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ
1				ļ	ľ	PVKHEDREIYPQHLKSLAFSAFTQCRRPLPTS
1		1		1		TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
i	ì	1	1	ļ	1	PPFILAPVKDKQTELGETFGEAGQKYNVLFV
ì	1	}				
1			ł			GYCLSHDQRWILASCTDLYGELLETCIINIDVP
		1	1			NRARRKKSSARKFGLQKLWEWCLGLVQMSS
		1			1	LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ
	1	1			1	SLSKRLKDMCRMCGISAADSPSILSACLVAM
1	1	1	1	i	1	EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL
		1		I		NTPQDTSCTHILVFPTSASVQVASATYTTENL
1	1 .	ſ	1	i	1	DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
		1			1	NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD
1	1	i		1	1	RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
1				1		LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
1		1		i	•	QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN
1	1 '	1	1	I		ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI
1	1	1	1	i		MNML
789	2139	HA-	6359	<del>  1</del>	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
1 '09	2139	1^	0555	1 *		LPVGPLLRALATCHALSRLQDTPVGDPMDLK
1		1		1		MVESTGWVLEEEPAADSAFGTQVLAVMRPP
1	1	1	1	1	!	LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM
			1			
1	1	İ	1	ļ		SVVVAWPGATQPEAYVKGSPELVAGLCNPET
1			1		1	VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
1		1	1		1	SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
		1	1		1	QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
1			1		1	RGCGMVAPQEHLIIVHATHPERGQPASLEFLP
1	1			İ		MESPTAVNGVKDPDQAASYTVEPDPRSRHLA
	1			1	Ī	LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
1 .	1	1	1	-	1	EQKTELVCELQKLQYCVGMCGDGANDCGAL
				1	1	KAADVGISLSQAEASVVSPFTSSMASIECVPM
L						

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG VQLGGYFLTLAQPWFYPLNRTVAAPDNLPNY ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN NVPFLLASAL*SSVLVVLVSPGLLHGPLALR NTTDTGFKLLLVGLVTLNFVGGLHAGERARP VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW PPLPAGPLR
790	2140	A	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEEILEQADYLYESGETEKLYQLLTQYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/L/EKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FPPYEKALGYFHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN
791	2141		6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ *VTICPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTIGIKQIKQEPIYDLTSVPNLFTY\SS FNNIGQLAPGITIMTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK
792	2142	A	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL
793	2143	A	6446	3201	152	PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDT\QFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

				Day dies 1	70 - 21 - 4 1	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	Ì	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	la isoleucine, Kallysme, La Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		<b>i</b>	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/=possible nucleotide deletion, \=possible
	ļ	1		sequence	,	nucleotide insertion
	<del></del>					IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL
		1		ļ	ļ	LEYCMVTGYDWWDILLHVQPSMVQSLVEKL
	<b>{</b>	1 .	Ì	}	1	HEEYTRQTAALQQVLSTRILAMKASLCKLSP
	1	1		ŀ		CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT
1	1			Į		PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF
i					}	VLDMNTLQALQQLLQWVGDFVLYLLASLPN
	ĺ		Į	Ī	1	QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLC\G
ì		1				SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL
Į.		1		ļ	]	KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC
i			1		1	RDEGPASEPDEALVDECCLLPSQLLIPSLDWL
l		1	1	Ì		PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT
1	ł		!	i	İ	LOLDGLARAPGOPKIDHLRRLHLGACPTEEC
		1				KACTRCGCVTMLKSPNRTTAVKQWEQRWIK
	1	1	1	İ		NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA
1		1	1			APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG
ļ.	1	1	l		Ì	RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT
í			1	1	ŀ	
	<u> </u>	<u> </u>				PRSLDHLHPEDRP
794	_2144	Α	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
<u> </u>	,			<u> </u>		GSIEPRDLRLQ*AVITPL\TPAWVTQ
795	2145	A	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
	1				1	SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP
	ì			l	1	FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP
1			1	(		DHIHHHLCRLSLESSPLFHHRVLFCVPKQNVN
	1	1	1		.	STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL
1	1 -	1		ł	1	HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS
	(	ĺ	(			DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA
1				1		DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR
ı	1		1			PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM
	į	1				KKMNKLQNISEELQRNISLQLMSNMNISNKIR
1	ŀ					NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR
1			1			WIWHKDSCYFLSDDVQTWQESKMACAAQN
1	i		ĺ			ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE
Ĭ			1	į		DS/YSWYESG*YNQ\PSAWVIRNAPDLNNMY
1	1					CGYINRLYVQYYHCTYKQRMICEKMANPVQ
1		1	1			LGSTYFREA
797	2147	$\frac{1}{A}$	6507	+1	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH
191	214/	A	0307	1	001	APRAASAQLEERMRDPHPGMTLQEGDCRGS
Ì	1		1	1		QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ
1						ESALAKLLLTCCSALRPRATQARGSSRLLVAS
1					l	WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA
1	1	İ		}	j	AIWTGAVAVLAGAAAFIYEKRGGTYWALLR
1	1	[	1	1	Ì	TLLALAAFSTAIAALKLWNEDFRYGYSYYNS
ļ	1	]		1		
	}		1	1	1	ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM
		1	1	1	1	DMLKALFRTLQAMLLGVWILLLLASLTPLWL
1	1 .		1	1	1	/SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG
			1	1	Į.	STHASARSQVPRSAGEAAPHSRRPPGLLPHAP
			1	ļ	1	RAASAQLEERMRDPHPGMTLQEGDCRGSQT
1	1		1	1	1	VSLTMGTADSDEMAPEAPQHTHIDVHIHQES
1	1		1			ALAKLLLTCCSALRPRATQARGSSRLLVASW
1	1	ļ	1	]		VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI
		1	1			WTGAVAVLAGAAAFIYEKRGGTYWALLRTL
		1	1	l l		LALAAFSTAIAALKLWNEDFRYGYSYYNSAC
1	1	J	1	1		RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM
1		1	1	1		LKALFRTLQAMLLGVWILLLLASLTPLWLYC
}						WRMFPTKGVSP
798	21/0	A	6528	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP
178	2148	Α.	0328	1 312	2207	RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG
						LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG
	L	ı	1	t ·	1	I PIA A AGIATWITE ATT I ASSSEDLI I MALQ. QUOQ

				· · · · · · · · · · · · · · · · · · ·		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
nuci-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ	]		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	İ	1	peptide		/=possible nucleotide deletion, \=possible
		J·		sequence		nucleotide insertion
	<del>                                     </del>					EVFPEGTGLPLPHSDLPTSWCGHSLQCGSQSS
	1	}	1	1	ŀ	FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK
	1	1			İ	HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV
	1	1				MVRWELSSNGNPGRGVLGLGLGLGNKLRVV
	l	1				GQNLGL*HCVWVVWETGE*KRWRLQMGIE*
	}	ł	1	}	}	GVASRRO*VRNSVRGLVCHNSSAPPMYMGFF
	İ				1	SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL
	t					GPSLPQRQGREHIVVILAAPACAPFHDR*WEP
	1	1				REIRPSP*ELGLRGEPTLSYPASCRVIRQPIP*D
Į.	1	1	1		ļ	RKSYSWKORLFIINFISFFSALAVYFRHNMYC
			1	ļ	1	EAGVYTIFAILEYTVVLTNMAFHMTAWWDF
ì		1	1	1	1	
					0.004	GNKELLITSQPEEKRF
799	2149	Α	6529	1	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
í	1	1	1	1	1 '	CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV
i		1	1		1	DEARCKESQQEAQENLREDLCLESFAKDKIL
			1		1	QIIEGSEREHEETRTKQAALDGEPLGGGQLTA
	1	1	1			VHLHPSKEQQGQEGGERQRGARTHHWRGW
1	1	1	1			EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T
			İ			ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV
1	1	1	1			RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL
	1	ì		1	1	GLPGDPTGPVTHHAPPVSPTGASGQERRAEP
ļ			İ			GAVSYAHASATK
800	2150	A	6544	2	662	SAORWAAVAGRWGCRLLALLLLVPGPGGAS
000	2130	1 ^	0544	<b>-</b>	002	EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG
l	Į.	1	İ	l .		GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF
Į.				1		TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG
1	ł	İ	1	ì		EXTHLCFLVR/DRVSALTQMESACVSIHEALKS
<u> </u>			1	1	1	VIDYOTHFRLREAQGRSRAEDLNTRVAYWSV
	j	1	į.	1	•	GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
			1	<del></del>	1319	TPCMECIKGEGLREPQNLSGSQREPQTEGSM
801	2151	Α	6556	1	1319	DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF
1					ļ	
	1					KRIFLKRMPSIRESLKERGVDMARLGPEWSQP
1	1		i			MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP
	i	-				PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY
1	1	1.	1			HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL
1	l .		Į.	j		SQDIITVGGITVTQMFGEVTEMPALPFMLAEF
1	1	1	ł	l		DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED
Į.				1	1	VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE
1		1			1	GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE
		-				DGCLALVDTGASYISGSTSSIEKLMEALGAKE
1						KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT
	1			1		SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL
	ł					\ALGATF\IRKFYTEFDRGNNPHGFALAR
802	2152	A	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL
002	2132	"	550/	]		LAVVVLLALPVAWGQCNAPEW\LPFARPTNL
1	-	ĺ		1		TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS
1		1		1		VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG
1		1				IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW
1						DNETPICDRIPCGLPPTITNGDFISTNRENFHY
1				1		GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND
	*	1		1	1	
1					1	DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
ŀ		1		1	1	NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA
	j	1	j	}	j	LNKWEPELPSCSRVCQPPPDVLHAERTQRDK
					1	DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
1		i		1	1	DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
			1	1		NLQLGAKVDFVCDEGFQLKGSSASYCVLAG
	-		1	ĺ		MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP
			1	1		LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR
				1		CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL
				1		FAKLKTOTNASDFPIGTSLKYECRPEYYGRPF
,		1	1	1	1	1

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence			914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide	sequence	/=possible nucleotide deletion, \=possible
1			ļ	sequence	,	nucleotide insertion
				<del></del>		SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG
						MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI
						LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS
			•			TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
		Ì				PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV   ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP
1		ļ	ļ			RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
	1		ļ		•	ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS
	ļ					MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
		i				GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS
ŀ						YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG
ļ	-			`	)	RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD   LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
						QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP
ł		İ	ļ			EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP
			•	į		PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG
l			1			HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI
l	İ	}	l	] .	ļ	ANGDFISTNRENFHYGSVVTYRCNLGSRGRK
					-	VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP
		1		1		GFVMKGPRRVKCQALNKWEPELPSCSRVCQ
				•		PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY
			1			DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF
				Ì		LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL
	]	l		, ·	ļ	KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
	-					PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
				1	Ì	GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS
			Ì			LNYECRPGYFGKMFSISCLENLVWSSVEDNC
				· .	1	RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC
		1			]	NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII
					·	SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH
						TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
				ļ		PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI   IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH
						CSRVCQPPPEILHGEHTLSHODNFSPGQEVFY
					1	SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
		1	ļ		-	KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC
					ļ	DEGFRLKGRSASHCVLAGMKALWNSSVPVC
		1				EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT   CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS
		1		1		SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL
				1		YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS
1		ļ		1	1	QLDHYCKEVNCSFPLFMNGISKELEMKKVYH
						YGDYVTLKCEDGYTLEGSPWSQCQADDRWD
ĺ		1		· ·	1	PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI
	1				1	ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP
803	2153	A	6574	2	3233	RTLQTNEENSRVLP HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL
003	2133	1^	05/4	1	1233	LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY
				{	1	PWSWA\RVGPAVELALAQVKARPDLLPGWT
1	1		1	1	1	VRTVLGSSENALGVCSDTAAPLAAVDLKWE
1						HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL
		1		1	1	TAGAPALGFGVKDEYALTTRAGPSYAKLGDF
		i		1	,	VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT
	1	J	İ		1	RLLRTMPRKGRVIYICSSPDAFRTLMILLALEA
		1				GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
		1			!	ERGDGQDVSARQAFQAAKIITYKDPDNPEYL
	<u> </u>	<u> </u>	<u> </u>		<u> </u>	EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Bonec		914	ng to first	acid residue	
uchoc			714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
i		l .		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan,
<b>i</b> 1		<b>!</b>			seduence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/-possible nucleotide deletion, \-possible
<u> </u>				sequence		nucleotide insertion
i i					1	DGLLLYIQAVTETLAHGGTVTDGENITQRMW
						NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE
						NGAFRVVLNYNGTSQELVAVSGRKLNWPLG
1						YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV
i l						GSLSLLGILIVSFFIYRKMQLEKELASELWRVR
					·	WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL
1 1						LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
1 1					,	IELTRKVLFELKHMRDVQNEHLTRFVGACTD
1 1						PPNICILTEYCPRGSLQDILENESITLDWMFRY
!!!						SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
1 .						DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
1						LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
1						IALRSGVFHVEGLDLSPKEIIERVTRGEOPPFR
1 :						PSLALQSHLEELGLLMQRCWAEDPQERPPFQ
1		l				QIRLTLRKFNRENSSNILDNLLSRMEQYANNL
1						EELVEERTQAYLEEKRKAEALLYQILPHSVAE
						QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE
1 1						STPMQVVTLLNDLYTCFDAVIDNFDVYKVET
1		İ			,	IGDAYMVVSGLPVRNGRLHACEVARMALAL
1 1					,	LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV
1 1						VGLKMPRYCLFGDTVNTASRMESNGEAL\KI
						HLSS\ETKAVL\EEFGGFELELRGDVEMKGKG
						KVRTYWLLGERGSSTRG
804	2154	A	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
1				_		MSERVSGLAGSTYREFERLIVRYDEEVVKELIP
1					ě	LVVAVLENLDSVFAQDQEHQVELELLRDDNE
) !						QLITQYEREKALRKHAEEKFIEFEDSQEQEKK
						DLQTRVESLESQTRQLELKAKNYADQISILEE
1 1						REAELKKEYNALHQRHTEMIHNYMEHLERT
.[						KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP
[						AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
1						SHTSLKDELSDVSQGGSKATTPASTANSDVA
<b>!</b>						TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV
1 1			·	į		QVAQETRNVSTGSAENEEKSEVQAIIESTPEL
, ,						DMDKDLSGYKGSSTPTKGIENKAFDRNTESL
i						
			İ			FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK
1 1						
						DVLQGELEAVKQAKLKLEEKNRELEEELRKA
		l				RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE
] ]					]	MARVLMERNQYKERLMELQEAVRWTEMIR
						ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK
]						KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP
]					i	GDKSKAFDFLSEETEASLASRREQKREQYRQ
] [	. [			[		VKAHVQKEDGRVQAFGWSLPQKYKQVTNG
			ļ			QGENKMKNLPVPVYLRPLDEKDTSMKLWCA
]						VGVNLSGGKTRDGGSVVGASVFYKDVAGLD
[				į	ĺ	TEGSKQRSASQSSLDKLDQELKEQQKELKNQ
		1				EELSSLVWICTSTHSATKVLIIDAVQPGNILDS
[			.			FTVCNSHVLCIASVPGARETDYPAGEDLSESG
1		1			ľ	QVDKASLCGSMTSNSSAETDSLLGGITVVGC
		, 1		ł		SAEGVTGAATSPSTNGASPVMDKPPEMEAEN
				i		SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS
				ł		QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND
				1	. [	SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
						MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD
}	1	1		ŀ	į	SILSIVHVKGIVLVALADGTLAIFHRGVDGOW
				1	İ	DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG
[				İ	1	YRNKIYVVQPKAMKIEKSFDAHPRKESOVRO
			1	l		LAWVGDGVWVSIRLDSTLRLYHAHTYOHLO
L		i				DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ucinos	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciico		Į	7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			]	peptide		/=possible nucleotide deletion, \=possible
		ł		sequence		nucleotide insertion
			<del> </del>	Soquetion	· · · · · · · · · · · · · · · · · · ·	NRLWVGTGNGVIISIPLTETVILHQGRLLGLR
				l		ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT
		i	1	1		FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV
			ŀ		]	ISPOSSSSGTDLTGDKGRGHLHRSLVVRRP
006	2155	A	6605	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ
805	2133	^	0003	رس ا	1 2002	SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP
	1	ì	Į			VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS
		1	į.	1		SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ
		ì			}	DSGLYACVIRNSTYCMKVSISLTVGENDTGL
			Ĭ			CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
		ì		Į	<b>}</b>	REPEILWYKECRTKTWRPSIVFKRDTLLIREV
		1	Į.			REDDIGNYTCELKYGGFVVRRTTELTVTAPL
	1	ì	ĺ	1		TDKPPKLLYPMESKLTIQETQLGDSANLTCRA
İ	İ	1	ļ	1		FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE
ļ		[		1		SDINKILKEHLGEQEVSISLIVDSVEEGDLGNYS
		ì	į	Ì		CYVENGNGRRHASVLLHKRELMYTVELAGG
		1	1	1		LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA
		1		1	1	EELDGDNKDYDAYLSYTKVDPDQWNQETGE
	<b>,</b>	1.	1			EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT
1	1	1		İ	1	YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF
	J	į.	1	1	1	ELETRLRNMLVTGEIKVILIECSELRGIMNYQE
ĺ	1	1	1		1	
		1			1	VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR
		1		1		LQYEMPFKRIEPITHEQALDVSEQGPFGELQT
	1	1	ì	1	}	VSAISMAAATSTALATAHPDLRSTFHNTYHS
	}		]	1	1	QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT
1	1				1	YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA
				<u></u> .		ILPLLPRETSISSVIW
806	2156	Α	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL
1	1	1				HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT
	j		ļ	1		QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL
	1.	1	1	1		IQNKYFGDVDIPRAKVVRVCQALMDYKVFE
	1	1	1	<b>\</b>	-	AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD
· ·		1			1	SQLGKENKLYSPARYADALFKSSDIRSASLED
1		l .	i	1	1	LWENLSLKPANSPHVNISTTLSPQVINEVWQE
		ļ	1	1		ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK
į.	ĺ	-	1		ì	ROSTMVNSSNYLDRGILKAYSDSQEDEWLSA
1				1		AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE
l		1				LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
		1	l l			NGKTEIALEATQLLLKLLDFQNREEFRRLLYF
1	1	}				MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
1		4				DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT
	1	[			J	1 - 12 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 -
1				ı	i i	L\HKIVS\VK\LMAIQNGRDPNRDAGYTYCQRI
					ļ	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA
						DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR
807	2157	Ā	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL OPGAFQALTHLEHLSLAHNRLAMATALSAG
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL OPGAFQALTHLEHLSLAHNRLAMATALSAG
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD  FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKMVKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD  FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKMVKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD  FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKMVKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLLALLTLGLAAQHQDKYPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE TLELGARALG\SLRTLLLQGNALRDLPPYTFA
807	2157	A	6615	4198	2094	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD  FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKMVKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LDLSSNPGLEVATGALGGLEASLEVLALQGN GLMVLQVDLPCFICLKRLNLAENRLSHLPAW TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI NLIIILTFILVSAILLTTLAACCCVRRQKFNQQ YKA
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG ARHRNNTEKKHPGGESDASPEAGSGGGGV ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN AGSVGLALIVWIVTGFITVVGALCYAELGVNI PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA AICLLLTWVNCSSVRWATRVQDIFTAGKLL ALALIIIMGIVQICKGEYFWLEPKNAFENFQEP DIGLVALAFLQGSFAYGGWNFLNYVTEELV DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT AMSPQEL\LAS\NAVAVTFGEKLLGVMAWIM PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD MYTLINYVGFINYLFYGVTVAGQIVLRWKKP DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI ELLTLVSQKMCVVVYPEVERGSGTEEANED MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY LFTSPQRVYRNFHYRKQTKDQWARDDPAFL VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV VLIDCVGVGLLIATLMWFISNKYLVKRQSRD YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL GYSVGLLFFSVALPFLKNTVILLYPFAPLILLYG LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811	2161	A	6627	18	3367	LEGSLNTERAK YYLTITMPHFT VTK VEDPEEG AAASISQEPSLADIKARIQDSDEPDLSQNSITG EHSQLLDDGHKKARNAYLNISNYEEGDEYF DKNLALFEEMDTRPK VSSLLNRMANYTNLT QGAKEHEEAENITEGKKKPTKTPQMGTFMG VYLPCLQNIFG VILFLRLTWV VGTAG VLQAF AIVLICCCCTMLTAISMSAIATNG VPAGGSY FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL GAIEIFL VYIVPRAAIFHSDDALKESAAMLINI MRVYGTAFL VLMVLVVFIG VRYVNKFASLFL ACVIVSILAIYAGAIKSSFAPHFPVCMLGNRT LSSRHID VCSKTKEINNMT VPSKLWGFFCNSS QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN LWSNYLPKGEIIEKPSAKSSD VLGSLNHEYVL VDITTSFTLL VGIFFPS VTGIMAGSNRSGDLKD AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV

PCT/US01/03800

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide seq- uence	peptide seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of peptide	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion  FGHSKANGEPTWALLLTAAIAELGILIASLDL VAPILSMFFLMCYLFVNLACALQTLLRTPNW RPRFRYYHWALSFMGMSICLALMFISSWYYA IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS LSAARFALLRLEEGPPHTKNWRPQLLVLLKL DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ LVVAAKLREGISHLIQSCGLGGMKHNTVVM GWPNGWRQSEDARAWKTFIGTVRVTTAAHL ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG GMLMLLPFLLK\QHKVWRKCSIRFTVAQLE DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS DISAYTYERTLMMEQRSQMLRHMRLSKTER DREAQLVKDRNSMLRLTSIGSDEDEETETYQ EKVHMTWTKDKYMASRGQKAKSMEGFQDL LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL ERVLLVRGGGSEVITIYS
812	2162	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ VAGADFESEDEGEEFDDWEDDYDYPEEEQLS GAGYRVSAALEEADKMFLRTREPALDGGFQ MHYEKTPFDQLAFIEELFSLMVVNRLTEELG CDEIIDRE
813	2163	A	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC WQYRQLSALHRAPRPTRPDKARRLGYKAKQ GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG ATYGKPVHHGVNQLKFARSLQSVAEERAGR HCGALRVLNSYWVGEDSTYKFFEVILIDFFHK AIRRNPDTQWITKPVHKHREMRGLTSAGRKS RGLGKGHKFHHTIGGSRRAAWRRNTLQLH RYR
814	2164	A	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN PLTKESIRQKEMESKRLRLLQEEDRRKKIARM GFNASSMLRKSQLGFLNVTNYCHLAHELRLS CMERKKVQIRSMDPSALASDRFNLILADTNS DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF MHENLYFTNRKVNSVCWASLNHLDSHILLC LMGLAETPGCATLLPASLFVNSHPAGIDRPG\ MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR RVLLTNVTGHRQSFGTNSDVLAQQFALMA PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF HDSAVTSVRILQDEQYLMASDMAGKIKLWD LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTV DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

PCT/US01/03800

SEO ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D-Aspartic Acid, E-Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	0000	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 552.55	j		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Į.	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]		ľ		peptide		/-possible nucleotide deletion, \-possible
ĺ	]		ļ	sequence		nucleotide insertion
						RVFETLKDLKVLNLAYNKINKIADEAFYGLD
1	1		1		ļ	NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL
1				ļ		QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH
1		1	Ì			FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR
1		1			1	LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ
l	i	ł		1	ł	TPSENPSLEQLFLGENMLQLAWETELCWDVF
		İ	1			EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR
1		i			1	GLSLNSNRLTVLSHNDLPANLEILDISRNQLL
		1		1		APNPDVFVSLSVLDITHNKFICECELSTFINWL
}	}	}		]	1	NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE
	1	Į.				GCDEEEVLKSLKFSLFIVCTVTLTLFLMTILTV
Ì	l,		1	1	1	TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY
		1				KYDAYLCFSSKDFTWVQNALLKHLDTQYSD
ļ.		1	1	İ		QNRFNLCFEERDFVPGENRP\ANIQDAIWNSR
1		{	Ì	1	1	KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL
		1	1		l	NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ
1		1				YLRWPEDLQDVGWFLHKLSQQILKKEKEKK
		1			-	KDNNIPLQTVATIS
816	2166	A	6646	1	3811	RDRAGVRPAGKQHAAAAFYDVGGDRPWDS
		1				GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
	ĺ			1		GGARLASLFGLDQAAAGHGNEFFQYTAPKQP
1		į		i		KKGQGTAATGNQATPKTAPATMSTPTILVAT
j	}			İ	1	AVHAYRYTNGQYVKQGKFGAAVLGNHITR
	l				-	EYRILLYISQQQPVTVARIHVNFELMVRPNNY
1	· .		1 .	.0	1	STFYDDQRQNWSIMFESEKAAVEFNKQVCIA
		1	1	İ		KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE
İ		ļ		1	1 .	VAYTGWLFQNHVLGQVFDSTANKDKLLRLK
1	1	1	١,	i	1	LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA
				1	]	CAVGSEGVIGWTQATDSILVFEVEVRRVKIA
		1	1	1	1	KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV
				1		VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD
		1	1			AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI
			1	1	İ	EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ
1		1		1		MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA
	İ					VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ
	1	1		]	}	PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR
1	1	1	ł			QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
1	•	1	-			LKQEILEKSNRIEEQNDKISELIERNQRYVEQS
1		1	1	1		NLMMEKRNNSLQTATENTQARVLHAEQEKA
-			ł	İ	•	KVTEELAAATAQVSHLQLKMTAHQKKETEL
1	1	1				QMQLTESLKETDLLRGQLTKVQAKLSELQET
1				1		SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL
		1	1	l		RVEKESLEKNLSERKKKSAQERSQAEEEIDEI
		1		ļ		RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS
}	ľ	1	Ì		1	LVQAELQTQWEAKCEHLLASAKDEHLQQYQ
			1	1	1	EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA
				1	1	OITALTKONEOHIKELEKNKSOMSGVEAAAS
i		1		1	<b>\</b>	DPSEKVKKIMNQVFQSLRREFELEESYNGRTI
}	1			1		LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE
-	1		i	1	1	EKAEERPRRPSQEQSASASSGQPQAPLNRERP
		1		1		ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR
1			-	1		KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP
	1					TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN
	1	1	ĺ			PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE
		-	1	1		GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
	1	1	}			ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE
ĺ		1				DELFKGATLKALRPKAQPEEEDEDEVSMKGR
	1	- 1	- 1	1	1	PPPTPLFGDDDDDDDDDDWLG
817	2167	A	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG
01/	2107		7047			

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIODLMDIFG DRYCALNNKATGAEQEAQRAQLLGLIQRVV RENKEGCYTNRMYQRAEEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFFYTDNLWDTLPCSWQEALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPGYHLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCEELLLPLENPCQGRA RLLLTGLHACGYDLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRPGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSPYGFHAELLPIFSPELSPRNLVLVATKMPLG QALSVLETEDS
819	2169	A	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGTILDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARNDMRSCLTEGVEAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAGVRGNVEKKPSSVRALSIVL PIVLLVFLCLGVFILWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \*-possible nucleotide insertion
						EMTNLKDIGLYNLRNITRGAIRIEKNADLCYL STVDWSLILDAVSNNYIVGNKPPKECGDLCP GTMEEKPMCEKTTINNEYNYRCWTTNRCQK MCPSTCGKRACTENNECCHPECLGSCSAPDN DTACVACRHYYYAGVCVPACPPNTYRFEGW RCVDRDFCANILSAESSDSEGFVIHDGECMQE CPSGFIRNGSQSMYCIPCEGPCPKVCEEEKKT KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFLK NLRLILGEEQLEGNYSFYVLDNQNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHF TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFKNVTEYDGQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPERKRRDVMQVA NTTMSSRSRNTTAADTYNITDPEELETEYPFF ESRVDNKERTVISNLRPFTLYRIDHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWELATLAEQPYQ QLSNEQVLRFVNMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF
821	2171	A	6691	106	825	DERQPYAHMNGGRKNERALPLPQSSTC GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEBIKTLSKEEETKK
822	2172	A	6715	772	21	DFRPGLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGD\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823	2173	Ą	6727	3	4063	PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS VGGQNPSTGGISADRTQGNIGCGGDTDPGQS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SSQPSQDGQESNVPSVGSLADPDYLNTPQMN TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL SDSVMNIFKDRNFDSCCICACNMNIKGADVG LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL FLEDELDIFGKNSDIGQAAERRLMMCQSTFL PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN FLDYISSNNRQTLPCVSWSYDRVQADNNDY WTECPNALEQGRQYVDNPTGGKVDEALVRS ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP
-						FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWFNQPW SGEENDNHSRLKLYAQVCRHHLAPYLATLQL DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM
						VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRPLPTQIHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL
824	2174	A	6732	2440	365	VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGGG
						GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGVKLESLSILIS SGVANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE
						DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD  G RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARA\TDYLQ ASAITRIPSYRYRYQRRSRSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA
826	2176	A	6744	3	5177	NRRTTPV  SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQA WQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEPPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYYDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTIL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGRQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

PCT/US01/03800

MO: of mole celide edide sequence peptide celide uSSN of 9496 914 914 914 914 914 914 914 914 914 914	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
neclecidide seq- uence  1914  1914  1914  1914  1914  1914  1916  1914  1916  1914  1916  1914  1916  1914  1916  1914  1916  1914  1916  1914  1916  1916  1916  1916  1916  1916  1916  1916  1916  1916  1917  1917  1918	•						D=Aspartic Acid, E=Glutamic Acid,
uence  USSN of 144  914  14 in the foresponding corresponding land in the state of pertide annino neid residue of pertide sequence  15 in the state of pertide sequence  15 in the state of pertide sequence  16 in the state of pertide sequence  17 in the state of pertide sequence  17 in the state of pertide sequence  18 in the state of pertide						location	F=Phenylalanine, G=Glycine, H=Histidine,
### BRY   14   15   16   16   16   16   16   16   16	_					corresponding	I=Isoleucine, K=Lysine, L=Leucine,
ueice  1914 a gin ins da minio acid de radiue flute minio acid de residue of peptide peptide peptide sequence 1915 peptide sequence 1916 a gin instantion acid of peptide sequence per per per per per per per per per pe				3			
amino acid residua of peptide residua of peptide residua of peptide sequence per control peptide sequence per control peptide sequence per control per	-	dellec	İ		•		
residue of peptide of peptide sequence per control peptide sequence peptide sequence per control peptide sequence per control	dence	ľ	İ	717			
peptide sequence    Poptide   Apposible nucleotide deletion,   Poposible     Poptide     Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide   Poptide   Poptide     Poptide   Poptide   Poptide   Poptide   Poptide   Poptide   Poptide     Poptide   P		ł	}				
827 2177 A 6748 2 1662 FIGURE SEQUENCE			1			Soquestion	
VOKKIMOVITKPIGGAAELVSQTGYGILHGA						1	· · · · · · · · · · · · · · · · · · ·
GLSQLFKQRHQPSDVHADQAPNSFIVKYVW KMLQSLGRFEVMMADDVVLNGSGGEGEG LLLTSEVLFVVSVSEDTQQQAFFVTEIDCAQD SQNNLLTVQLKQPKACDVEVDGVPGRBLSE QQYNRLDVTIKTSCHLAFSCSSMQDCPVVA AEPPETVKTYHTLVDHFAQVFLSKFTMVK NKALRKGFP FGAPRGNPFGSFGNPGRHQGPCHPRRGTK ASGVSPTLWRPQAAATGLEMPSGRALLDSF VGGARPVLSQQALSCPHTRLGKLAVVVASY RRFGALAAATGLEMPSGRALLDSF RRFGALAAVSSEGGEPLAGDCFTVN VGGSRFVLSQQALSCPHTRLGKLAVVVASY RRFGALAAVSPELCDDANPVDMEYFERS SQAFRVLHYYNGRLIVMEQLCALSFLQBI QYWGDELSIDSCKDRYFRRKELSETLDFKK DTEDQESQUESEQDPSQGPCFTVRKLWML EKPOSSTAARFOVISHTVORSUMALMSAEL SWLDLQLLELLEYVGCLLALRALWALKAWG HSTGLRSJGMSTTTTVCEBVGLLLLFSVGISH TOPLENVGAHCPGCLRLALWALKAWG HSTGLRSJGMSTTTTVCEBVGLLLLFLSVGISH ANDRSFSACYFTLSKAAAVRQREAKKLTK NIATDSYTSVNLADVYARSIMEMILRLKGRER ASTRSSGODFVP ASSGGOPTAPSGGPTTRSVCAWWATTSMT TVGYCDIRPDTTTGKLVAFMCLSGILVLAUP ANDRSFSACYFTLSKAAAVRQREAKKLTK NIATDSYTSVNLADVYARSIMEMILRLKGRER ASTRSSGODFVP ASSGGOPTOSKGGHTFRALPGPGSMGLK NGAGNGAKGKRERSISADSFORQDGTOTN DSDBECNSADHIKSVSSPSGNTGSSKSKQ EMWYRPTTVNSPSGRPQLDSKFSNQGQGGS ASQSGPSCBKSGGHTFRALPGPGSMGLK NGAGNGAKGKGRRSISADSFORQDFGTTN DDSDBECNSADHIKSDQSGTTGPHINSNNK TERSTAPLNTQISALRNDPKRI-QQFPAPANQ DONSSQNTLQPTPPPAPARKAPPIPLDEE SGVENKLIPSVGSPASSTLUPDGTGPNSTTN NRAVTPYSQGSNISSASDPAAPPPPYSGGPPAPAD DEBEFTGAQSGGPQORDVLDCPGKFPGFGGHG DYFTSPDGMYPSISADSGGTGDPHDLGHTTP EQQAMLKLQQEFYEKRRFEQVYVQCCSLQ DMWHQRGFRGVVXGCPPFYQGTPSCWAP GGTEPSFDGNAPPISTSCWAP GGTEPSFDGNAPPISTPYCGGRERP NPQGLSEGMPQQLAEKQLGIPGGAMMGGR REMNAMGGSGGTGPHDLGHTTP FGQAMLKLQQEFYEKRRFEQVYVQCCSLQ DMWHQRGFRGVVXGCPPFYQGTSSCWAPPIPVEGP LSPSRGDFFKGPPQGGRSSGCRGRRF NPQGLSEGMPQQLAEKQLGIPGGAMMGGR KGDVNLNVMGGSGSGGTGPHTVPVGGC SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQQ SPMLUSPSGORNSLEFTNHOQGSUSSUSHSTHFPPFOTTS PEPLQSFGPPNIKAPPLTNSSSPFTSLNTAPPVCRG GPPPTASQCSASSSTANHASPSLAPPPPPT					sequence	<del> </del>	
KMLQSLGRFEVHMALDVVLVRGSGGEHEGG   LLITSEVLFVVSVSEDTOQQAFFVEIDCAQD    SKQNNLLTVQLKQPRVACDVEVDGVRBRLSE    QQYNRLDVTIKTSCHLAFSCSSMQDCPVVA    AEPPSTVKTYHTLVDPHFAQVFLSKFTMVK    WKALRKGFP    FGAPRRGSPFGSFGNPGRHQGPCHPRGTK    ASGVSPTLWRQAAATGLEAPPSGRALDSP    LDSGSLTSLDSSVFCSEGGEPLALGDCFVVN    VGGSRFVLSQQALSCPFHTRLGKLAVVVASY    RRFGALAAVFSTELLDDANNVONBFFDRS    SQARRYVLHYYSTGRLHVMEQLCALSFLQEI    QVWGDISLIDSCKORFYFRKSLESTLDFKK    DTEDQESQHESSQDFSQGPCTVRQKLWNIL    EEPOSSTAARFGVSIENTYVSINMALMSAEL    SWLDLQLLELLEYVCISWFTGEFVLRFLCVRD    CRETICAL AVFSTELLLIFYTILLVSSGSQT    TOELSINVGAHCPGCLRLRALIRMLKAWGR    RKTELSGAMTTQCVEVGLLLLIFYTILLVSSLGSQT    TOELSINVGAHCPGCLRLRALIRMLKAWGR    KTGLRSLGMTTQCVEVGLLLLIFYTILLVSLSGSQT    TOELSINVGAHCPGCLRLRALIRMLKAWGR    KTGLRSGMTVLFYVARSMEMLRLKGGER    ASTRSSGGDDFWP    ANDRESSGVFTLKLKRAAVRGREALKKLTK    AINDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ANDRFSACYFTLKLKRAAVRGREALKKLTK    ASTRSSGGDDFWP    ASTRSSGGDDFWP    ASTRSSGGDDFWP    ASTRSSGGDDFWP    ASTRSSGGDDFWP    TASSGGDSMTSNSCRACHRONLSNSK    TERTARALTOLISALRNDRKRLPQPFPAPARNAPHTPSACT    ARNOTYPYSQGSNSSSKANSDSGRTGPHNSTNSAT    ARNOTYPYSQGSNSSSKANSDSGRTGPHNSTNSAT    ARNOTYPYSQGSNSSSKANSDSGRTGPHNSTNSAT    ARNOTYPYSQGSNSSSKANSBGRTGPHNPTPSGGAANGSGRTGPHNSTSGWAP   ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGRTTPNAPCKGA    ANDRFSGGAANTSSEGGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANTSGGAANT		1	ļ	1	ļ	1	CI SOI DEODRIODEDIVILLADO ADMENVE VAN
\$27 2177 A 6748 2 1662 FUNDAMENTAL PROPERTY OF SERVING PROPERTY OF		ì		1	l		WALOU ODDENIA ALDIA MOSCOPHECO
SKONNLUTYIKISCHLAPSCSMOPCEVVA APPPSTVKIYHYLOPPIPAQVELSKFTMVK NALRKGPP  2 1662 FVGAPRGONFOSPONPORHOGOCHEPPERIC ASGVSPTIJWRQAATGLEMBSGRALLDSP LDSGSLTSLDSSVFCSEGEGEPLALGDCTVVA VOGSRYLSQQALSCFPHTRLGKLAVVVASY RPOGALAVPSPLELCDDANPVONBYFFDRS SQAFRYVLHYYRTGRLHVMEQLCALSFTQD QVWGIDELIDSCCRDRYFRRKELSETLDPKK DTEDQESQHESEQDESQGPCFTVRQKLWNI, EKPGSSTAALRGVSHIFVQSIBMALMSABL SWLDLQLLELLEYVCISWFTGEFVLRELCVRD RCRFLRKVPHIBLLALRFYHTLYSLGSGST TOELJENVGAHCPGCLRLLALRAMLAWMA HSTGLRSLGMTTDCYCEVGLLLLESJGSGST TOELJENVGAHCPGCLRLLALRAMLAWMIA KINATDSYISVILRDTYGYEAWWWATTSMI AINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVILRDTYARSIMBHLKGRER ASTRSSGGDFWF TASSGPPCDSKSGGHTFSVPCAWWATTSMI NIATDSYISVILRDTYARSIMBHLKGRER ASTRSSGGDFW  528 2178 A 6786 5672 1360 GTHPASSGPYDEPAAVSAATREELGEPVPFV TASSGPPCDSKSGGHTFSSPSKOKGOGGS ASQSPSPCDSKSGGHTFSSPSKOKGOGGS ASQSPSPCDSKSGGHTFKALPCPGGSMGLK NGAGNGAKGKGRRERSISADSFDQRDFGTPN DDSDKECNSADBHSSQDGHTPHSWTPSNAT APRSSTPPHQTTATEPTAQKTTAKVVYYS TEMANKAABAVKGQVETIVSHMOKSNK TERSTAPLNTQISALRRDFRFLPQGPFAFAN DONSSONTRLQFTPPTPARAFFRAPPRFLDRE SPGVENKLIPSVGSSADSTFLPPDGTGPNSTY NRAVTPVSQGSNSSADSKAPPPVSSGEPPT LGENPDGLSQEQLEHERISLQTLRDIQRMLT DEFTGAQSGGPQNGFVLDOPQKFEGEPI QAMMAQSGLGKGFPFBTWASGFPPOGGREGPT NRAVTPVSQGSNSSADSKAPPPVSSGEPPT LGENPDGLSQEQLEHERISLQTLRDIQRMLT DEFTGAQSGGPQNGFVLDOPQKFEGEPI QAMMAQSGLGKGFPFBTWASGFPQGHTSGWA GGTEPPSDGNMPHISLPPRGMAPPPNMFSGWA GGTEPPSDGNMPHSLPPRGGBFSGWA GGTEPPSDGNMPHSLPPRGGBFSGWA GGTEPPSDGNMPHSLPPRGGBFSGWA GGTEPPSDGNMPHSLPPRGGFSGWA GGTEPPSDGNMPHSLPPRGMAPPPNMFGG LSRSRGPPFRGIPPQGGRSGFGRAGAGFBE MLKLRFGGSDMLPAQGKWFLFFFGLFGQD GRTSPSDGNMPHSLPPRGRAFGFODD RTNSRLSFMPHLLENSSNTISLNTAPPVGG LSRSRGPPFKGIPPQMGPGRELEFGMVPFSGM KGDVNLNVMNGSGSMFTGRREGFGPDD RTNSRLSFMPHLLNSSNTTSLNTAPPVGG GTEPPSDGNPNIKAPLTNASSPTISLNTAPPTQG GTEPPSGGPPTIGKSPTTMPFEPTL SPRMMRMFGSQNHMEPGLNSSPTTMPFEPTL GPPPLASSGFPNIKAPLTNASSPTSLNTAPPTQG GTEPPSGGPPTIGKSPTTMPFEPTL GPPPLASSGFPNIKAPLTNASSPTSLNTAPPTQG GTEPPSGGPPTIGKSPTTMPFEPTL GPPPLASSGFPNIKAPLTNASSPTSLNTAPPTQGF SPMMRMFTGSQNHMERFGFODD RTNSRLSFMPHLLSPSSNTTSLTPTHPPEPTL GPPPLASSGFPNIKAPLTNASSPTSLTMPPTCGFF GPPPTASO		Į.	ľ	ľ	ľ	Ţ	KINTOSTORIE AHMATDA ATA KOROOGEHEGE
827 2177 A 6748 2 1662 FVGAPRRONPGSPONPORNOGPCHEPROTY AAPPESTVKTYHYLVDPHAQVYLSKFTMVK NKALRKGFP FVGAPRRONPFGSPONPORNOGPCHEPROTY AGVSPTILWRQAAATGLEMPSSGRALLDSP LDSGSLTSLDSSVFCSEGGEBPLALGDCFTVN VGGSRFVLSQQALSCPPHTRIGKLAVVVASY RRPGALAAVPSPLELCDDANFVDNEYFFDRS SQAPRAVNHYYRTGRIHVMEGLCALSFLQEI QYWGIDELSIDSCCCDRYFRRKELSETLDFK DTEDQESQHESSEGDFSQCPCPTVRQKLWNIL EKPGSSTAARIFGVISHIFVGVSINMAALMSAEL SWLDLQLLELRYVISHVFGGVSINMAALMSAEL SWLDLQLLELRYVISHVFGGVSINMAALMSAEL SWLDLQLLELRYVISHVFGGVSINMAALMSAEL SWLDLQLLELRYVISHVFGGVSINMAALMSAEL SWLDLQLLELRYVISHVFGGVSINMAALMSAEL SWLDLQLLELRYVISHVFGGFVAGV RCFFLRKYPNIDLLAHPYTITLLVSSLSGGVS RCFFLRKYPNIDLLAHPYTITLLVSSLSGGVS RCFFLRKYPNIDLLAHPYTITLLVSSLSGGVS RCFFLRKYPNIDLLAHPYTITLLVSSLSGGVS RCFFLRKYPNIDLLAHPYTITLLVSSLSGGVS STVEYFAGSSBFDTTFTSVPCAWWATTSINF TVQYDDRPDTTTGKTVAAFMCILSGILVLAH AINDRYSAVYPILKREAVYQARSHEMALKGGRR ASTRSSGGDFFWTFTSVPCAWWATTSINF TVQYDDRPDTTTGKTVAAFMCILSGILVLAH AINDRYSAVYPLARSBFRMARLKGRER ASTRSSGGDFFWTFTSVPCAWWATTSINF TVQYDDRPDTTTGKTVAAFMCILSGILVLAH AINDRYSAVYPLARSBFRONQDSSFSNQCKQGGS ASQSQSSCDSKSGGHTFKALPGGGSNGCKGRE ASTRSSGGDFFWFTFTAAVSAATREELGEPVPFV TASSGFVDSWHSSNPKVRSSPSGNTQSSPKSRSQCKQGGS ASQSQSSCDSKSGGHTFKALPGGGSDRGFGFN TASSGFCDSWHSSNPKVRSSPSGNCKQGGG ASQSQSSCDSKSGGHTFKALPGFGSNDGLK NGAGNAKCKGKRERISADSPROPGTN DDSDKLECNSADHIRSQDSGHTPHSMTPSSNC APRESTSPPHQUTTATEPTPAAKTPAAVYVYVS TEMANKAABAVLKGQVETIVSFHQNISNNK TERSTAPINTQISALRNDFERIPGPRAFFF NGAGLARRDFKALPGFPAAPPPLDER SPGVEMLIFSVGSPASSTPLEPGPAAPPRLDBE SPGVEMLIFSVGSSASSTPLPGPGNSTFTN NRAVTPVSQGSNSSADPKAPPPPVSSGEPFT LGENPGLSQEQLEHRERSLQTLRQURMLFP DEKEFTGAQSGGGQQNGFGLDDGPQKKPEGPI QAMMAQSSLGKGFGFRTDVGAFFGFGMVSG MRLPGFAGMINSGSGRMEGFGNNPTPRBVEGP LSSSGGTPRSDGNMPHISLPPRGMAPPHNMOSQ MRLPGFAGMINSGGRMEGFGNNPTPRBVEGP LSSSGGPPRHENSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG				Į	<u> </u>	Į.	
### APPPSTVKTYHYLVDPHFAQVFLSKFTMVK NALRKGFP  ### STAPLING    #	1	ł	1	ļ	ļ		SKQNNLLTVQLKQPRVACDVEVDGVRERLSE
827 2177 A 6748 2 1662 FVGAPRRGNPGSPGNPGREQOPCHERRGTK ASGVSPTLWRPQAAATGLEAPSSGRALLDSP LDSGSLTSLDSSVTCSEGEGEPLALGDCTTVN VGGSRFVLSQQALSCPHTRLGKLAVVVASY RRFGALAAVPSPLELCDDANFVDNSYFPDSS SQAFRYVLHYYSTGRLHWRQLCALSFLQEI QYWGIDELSIDSCCRDRYFRRKELSETLDFKK DTEDQESQHESQDFSQGPCTVRQKLWNIL EKFGSSTAARIFGVISIIFVGVSIDMALMSAEL SWLDLQLLEILEYVGISWFTGEFVLRECVEND RCFFLRKVPNIIDLLAILFYYTLLVESLGSGYT TQELENVGARCPGCLRLLARIAMLKAWGR HSTGLRSLGMTTTQCYEEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFEVGLLLLFLSVGISF STVEYFAGSIBPTTTSVFCAFLAKIRMLKAWGR HSTGLRSLGMTTQCYFTLKLEAAVFACHAGARGA ASTRSSGODFWF TLASSGGDASGRALKKIRATSSASTANGSPSKAK GENGAGGAS ASGSGSPCDSKSGGHTPKALGGPGVGAGGAS ASGSGSPCDSKSGGHTPKALGFGGGGGGAGAS ASGSGSPCDSKSGGHTPKALGFGGGGGLK NGAGGAGAKGKRERSISASSFDQRDFGTFTN DDSDIKECNSADHIKSQDSQHTPHSMTTSNAT APRSSTPPHGGTAALFTAVTVYVFS TEMANKAABAALKGQVETIVSFHQNISNINK TERSTAPLNTGISALARIDFKTAVVYVYFS TEMANKAABAALKGQVETIVSFHQDRISNINK TERSTAPLNTGISALARIDFKTAVVYVYFS TEMANKAABAALKGQVETIVSFHQDRAFPALDRE SPGYENLLTSVGSAFGTAVTVYTS TEMANKAABAALKGQVETIVSFHQDRAFPALDRE SPGYENLLTSVGSAFGTAVTVYTS TEMANKAABAALKGQVETIVSFHQDRAFPALDRE SPGYENLLTSVGSAFGGTGPDILDDIMTT EQIAWLKLQQETYEKRRKFEQVVVQCCSLQ DMYNHQHGFRGVVRGPTPYQMTTSEGWF GGTEPSDGINMPHSLPPRGMAPHRNMCSQ MFLFGFGGMINFSMEGPTVYQMTGSGGFF SWFDDVKTPDGRNFTPQGGTFSGGRGFT NPQGLSEGFF NPQGLSEGFF PNPQGL		ŀ	Į.	İ		ļ	
827 2177 A 6748 2 1662 FVGAPRRGNPGSFGNFGRRQGFCRRPKT  A 6748 2 1662 FVGAPRRGNPGSFGNFGRRQGFCRRPKT  A 6748 2 1662 FVGAPRRGNPGAATGLEAPSGRALLDSP  LDGGLTSLDSSVFCSEGEGEPLALGDCTVN  VGGSRPVLSQQALSCFPHTRLGKLAVVVASY  RRPGALAAVPSRLELCDDAWPONEYFFDRS  SQAFRYVLHYYRTGRHVMEQLCALSFLQEI  QYWGIDE.SIDSCCRRYYRRKELSETLDFKK  DTDQGSOHESEQDFSQGPCPTVRQKLWNIL  EKRGSSTAARFGVISIIIVGVSINMALMSAEL  SWLDLQLLEILEYVCISWPTGEFVLRFLCVRD  RCFLRKVPRIDILLAILPFYILLVESLGSGYGT  TQELENVGAHCPGCLRLLRAURMLKAWGR  HSTGLRSIGMTTTCYEPSCHYRLLFSVGISF  STVEYFAEQSIPDTTFTSVPCAWWWATTSMT  TVGYGDIRPDTTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTFTSVPCAWWWATTSMT  TVGYGDIRPDTTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTFTSVPCAWWWATTSMT  TVGYGDIRPDTTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTFTSVPCAWWWATTSMT  TVGYGDIRPDTTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTFTSVPCAWWWATTSMT  TVGYGDIRPDTTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPDTTGKIVAFMCLSGILVLLFSVGISF  STVEYFAEQSIPTTGKIVAFMCLSGILVLLFSVGISF  TVGYFAEQSIPTTGKIVAFMCLSGILVLFSVGISF  TVGYGAMAGNAGKGKRERSISADSFORDPFPF  DDSDLECCNSADHTKSDGSOHTPHSMTPSNAT  APRSSTPPHQQTLATEPTPAQKTPAKVVYVYFS  TEMANKAAEAVLKQOYETIVSFHQISHNK  TEKSTAPLNTQISSALMPINETPROPTGMTSNK  TEKSTAPLNTQISSALMPNETPROPTGMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPNETPROPTGRMTSNK  TEKSTAPLNTQISSALMPROPTPPYQMTPSGMPD  DDSDLECCNSADHTHSMTSNSGGFT  DDDMTNGHGFRGVVRGPPPPYQMTPSGMPD  DDMNSQNTBLQTFTRAMFRAKAAFFULLSC  SPOUNKLIPSTSDGINMPHSLFPRGMAHPNMMPSQ  MRATTPVSQCSNISSADPKAPPPPVSGGPT  DEKEFTGAGSGGPQMPGVLDCPOKKREGEP  NRATTPVSQCSNISSADPKAPPPPVSGGPT  DEKEFTGAGSGGPQMPGVLDCPOKKREGEP  NRATTPVSQCSNISSADPKAPPPPVSGGPT  DEKEFTGAGSGGPQMPGVLDCPOKKREGEP  NRATTPVSQCSNISSADPKAPPPPVSGGPT  DEKEFTGAGSGGGPT  RAMACASAPVHLLSSGGAT  RAMACASAPVHLLSSGALMPC  RAMACASAPVHLLSSA	İ	ł	1	1		1	-
ASGVSPTLWRPQAAATGLEMPSSGRALLDSP LDSGSLTSLDSSVTGEGEGEBPLALGDCTTVN VGGSRTVLSQQALSCYPHTRLGKLAVVASY RRPGALAAVPSPLELCDDAMPVDNEYFPDRS SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI QYWGIBLSIDSCCRRYPRRKELSETLDFKK DTEDGESQHESSQDFSQGPCTVRQKLWNIL EKFGSSTAARIFGVISIIPVGVSIDMALMSAEL SWLDLQLLEILELYVGISWFGEFVLRFLCVRD RCFFLRKVPNIIDLLAILFYYTLLVESLGSGOT TQELENVGARCPGCTLRLARIAMLKAWGR HSTGLRSLGMTTQCYEEVGLLLFLSVGISF STVEYFAEQSIPDTIFTSVPCAWWWATTSMT TVQYGDIRPDTITGKVARMCLSGILVLALPI AINDRFSACYFTLAKKAAVRQREALKKLTK NIATDSYISVNLADVYARSIBMEMLRLGGER ASTRSSGGDEWF TASSGGDDFWF TASSGGPSMHSSNIPKVRSSPSGNTOSSPKSKQ EVMVRPTVAMPSSPGNPQLDSKSSNGKQGGGS ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKKRRSISADSFDQRDPGTPN DDSDIKECNSADHKSQDSOHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAGKTFAKVVYVFS TEMANKAAEAVLKGVETVISYFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQQPPARNQ DQNSSQNTRLQTPTPPARAFKPAAPPRLDEE SPGVENKLIPSVGSPASSTPLPFDGTGRNSTPN NRAVTPVSQGSNSSADPKARPPRLOGE PGRENKLIPSVGSPASSTPLPFDGTGRNSTPN NRAVTPVSQGSNSSADPKARPPRLOGE SPGVENKLIPSVGSPASSTPLPFDGTGRNSTPN NRAVTPVSQGSNSSADPKARPPRLOGEP DEKEFTGAGSGGPQNPGVLDGPCKKFEGFI QAMMAQSGLKGFGFRIDVGAFFGPGGHR DVFFSPDEMYPPSSMFSQSTICDBILDIMTP EQLAVLKLQGEFYEEKRRKFRGVVVQCSLG DMMVHOHGFRGVVRGPPPPYQMTPSEGWAP GGTSPSDGDMPHSLPPRGMAPHRNMPGSQ MRLPGFAGMINSEMEGPNVPPNASSFGGRPP NPQGLSEEMFQQLARKGLGPFOMAHERIPVEGP LSPSRGDPFKGEPFQGAMPHENPMGSGRPP NPQGLSEEMFQQLARKGLGPFOMAHERIPVEGP LSPSRGDPFKGEPFQGGRPP NPQGLSEEMFQQQLARKGLGPFGHAMAE REMMRNINGSQRHMEPGNNPIRPFYEGP LSPSRGDPFKGEPFQMAPHFPRGD GLGRPPTLSNORMAPHPSNNFISLTNTAPPVQG GLGRPPTLSNORMAPHSNIPSSNFISLTNTAPPVGG LSRSPNLGSNAASPYHLLSSFLAPSSFDATSS GPPPTASSORAASPYHLLSSFIJARSPGWTSS GPPPTASSORAASPYHLSSFIJARSPGAWTSS GPPPTASSORAASPYHLSSFIJARSPGAWTSS GPPPTASSORAASPYHLSSFIJARSPGAMTSS GPPPTASSORAASPYHLSSFIJARSPGAWTSS GPPPTASSORAASPYHLSSFIJARSPGAWTSS GPPPTASSORAASPYHLGSFISTYTMPPEPGL FPPPLSSFIFTHNASPAANLGNVESG GPPPTASSORAASPYHLGSFISTYTMPPEPGL FPPPLSSFIFTHNASPAANLGNVESG GPPPTASSORAASPYHLGSFISTYTMPPEPGL FPPLSSFIFTHNASPAANLGNVESG GPPPTASSORAASPYHLGSFISTYTMPPEPGL FPPLSSFIFTHNASPAANLGNVESG GPPPTASSORAASPYHLGSFISTYTMPPEPGL FPPLSSFIFTHNASPAANLGNVESG GPPPTASSORAASPYH		ļ	1	Ì		l	NKALRKGFP
ASGVSPTLWRPQAATGLEMPSSGRALDSP LDSGSLTSLDSSVTCSGGGEGEPLALGDCTTVN VGGSRFVLSQQALSCFPHTRLGKLAVVVASY RRPGALAAVPSPLELCDAMPUSFYPERS SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI QYWGIDELSIDSCCRDRYFRRKELSETLDFKK DTEDQESQHESEQDPSQGPCTVRQKLWNIL EKROSSTAARIFGVISIIPVGVSIINMALMASEL SWLDLQLLEILELYVGISWFTGEFVLRFLCVRD RCFFLRKVPNIIDLLAILFPYTILLVESLGSQT TQELLENVGARCPGCTLRLARIMLKAWGR HSTGLRSLGMTTQCYEEVGLLLFLSVGISF STVEYFAEQSIPDTTTTSVPCAWWATTSMT TVQYGDIRPDTTTGKIVAFMCILSGILVLALPI AINDRFSACYFTLKLKBAAVRQREALKKLTK NIATDSYISVNLBDVY ARSIMEMELKKGRER ASTRSSGGDFWF  TASSGGDFWF  ASGCGSMHSSNFKVRSSPSGNTOSSPKSQ EVMRPFTVMSPSGNPQLDSKSGNGKQGGGS ASQSQPSPCDSKSGGHTPKALPOPGGSMGLK NGAGNGAKGKCRERSISADSFOQRDFGTPN DDSDKECNSADHKSQDSONHTPHSMTPSNAT APRSSTPPHQCTTATEFTPACKTFAKVVYVS TEMANKAAEAVLKGVWEITVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQOPPAANQ DQNSQNTRLQFTPIPAAPKPAAPPRLDEE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPVSQGSNSSADPKARPARAPRPRLDRE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPVSQGSNSSADPKARPARAPRPRLDRE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPVSQGSNSSADPKARPARAPRPRLDRE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPVSQGSNSSADPKARPARAPRPRLDRE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPSVGGCSNSSADPKRPAPARPRLDRE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPSVGGCSNSSADPKRPAPARPRLDRE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPSVGGCSNSSADPKRPAPARPRLDEE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPSVGGCSNSSADPRAPARAPARPRLDEE SPGVENKLIPSVGSPASSTPLPFDGTGRSTPN NRAVTPSVGGCFYFEKRRKFRFQVVVQCSLQ DMMVHOHGPRGVVRGPPPPVSGGEPF NPQGLSEEMFQQCJCARKQLGIPPGMAMEGR PSMEMNRMFGSQRHMEPGNNPIRFPEGWAP GGTSPSDGMAPHSPPRGGMAPHPNMPGSQ MELPGFAGMINSEMEGPNVPRASSRGGGRP NPQGLSEEMFQQCJCARKQLGIPPGMAMEGR PSMEMNRMIFGSQRHMEPGNNPIRFPYEGP LSPSRGDPFKGEPFQAGGGREP NPQGLSEEMFQQQCARKGLFGFTGCERPP NPQGLSEEMFQQCGSNSGLRNLREPGIPPG MKLRPGGSDMLPAQGANAPHELSFTMYEGP MKLRPGGSDMLPAQGANAPHELSFTMYEGP MKLRPGGSDMLPAQGANAPHELSFTMASPAMLGAGGFE MKLRPGGSDMLPAQGANAPHILKSFTMHPQVQ SPMLGSPSGNLKSPQTPSQLAGMLARPAAAA SIKSSPVLGSAAASPYHLKSFSIAPSFWTISS GPPPTASOQPASVIPIGGSLFTMAPPPVAG GTSPSTPLAGSAAASPHALKSFSIAPSFWTISS GPPPTASOQPASVIPIGGSLFTMAPPPARAF GFPPTASOPASVIPIGGSLFTMASPAMLGNESG GPPP	827	2177	A	6748	2	1662	FVGAPRRGNPFGSPGNPGRHQGPCHRPRGTK
LDSGSLTSLDSSVFCSEGGEPLALGOCTTVN VGGSRFVLSQOALSCPHTRLGKLAVVVASY RRFGALAAVFSPLELCDDANPVDRSYFDRS SQAFRVLHLYVTGRLHWHGQLCALSFLQEI QYWGDELSIDSCCRDRYFRRXELSETLDFKK DTEDQSSQHESQDFSQPCPTVRQKLWNIL EKRGSSTAARIFGVISIIFVGVSINMALMSAEL SWLDLQLLEILEYVCISIWPTGEFVLRFLCVRD RCRFLRKVPNIDLLAILFFYTILLVESLSGSQT TQELIENVGAHCPGCLRLLRALRMLKAWGR RSTGLRSLGMTTQCYESVGLILFSLVGIST TVGYGDIRPDTTTGKVAPMCILSGILVLAIL AIDDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDDFWF ASTRSSGGDDFWF ASTRSSGGDDFWF ASTRSSGGDDFWF TASSGFQSMHSSDNSVKSSSSGNTQSSPFSKK EVMYRPTYMSPSGRPQLDSKFSNQGKGGGS ASQSSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKRERSISADSFDQRDPGTPN DDSDKECRSADHKSQDSOHTPHSMTPSNAT APRSTPPHGQTTATEPTPAQKTPAKVYVFS TEMANKAAEAVLKQVETIVSFHQNISNNK TERSTAPLNTQISALRNDPFRLPDGTGRNSTPN APRSTPPHGGTTATEPTPAQKTPAKVYVFS TEMANKAAEAVLKQVETIVSFHQNISNNK TERSTAPLNTQISALRNDPFRLPDGTGRNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSGGPT LGENPDGLSGGGPQONPGVLDGVCKKFEGGP QAMMAGSGLOKEFGFRTDVAPFGPQGHR DVPFSPDEMYPSSMSSGGTGPBLIDMTP EQUAWLKLQGEFYEEKRIKERGQVVVQCGSLQ DMMVHQHGPRGVVRGPPPPVQMTPSEGWAP GGTSPFSDGINMPHSLPPROMAPHNMPGSQ MELPGFAGMINSEMEGPNVPRPASRFGLSGWF NPQGLSEEMFQQQLAEKGERSTQTLRIDGRMLFP NPQGLSEEMFQQQLAEKGLQTLRIDGRMLFP EQUAWLKLQGEFYEEKRIKERFQVVVQCSLQ DMMVHQHGPRGVVRGPPPPVQMTPSEGWAP GGTSPFSDGINMPHSLPPROMAPHNMPGSQ MELPGFAGMINSEMEGPNVPRPASRFGLSGWF NPQGLSEEMFQQQLAEKGLGTPGMAMEGRI PSMEMNRMFGSQRHMEPONNFIFRIYEGP LSPRGDFKGEFFQNVPSGGBFT NPQGLSEEMFQQQLAEKGLGPTPVGGEP LSPRGDFKGEFFQNVPSGGBFT NPQGLSEEMFQQQLAEKGLGPTPVGGEP LSPRGDFKGEFFQNVPSGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ YGMGPRFFLPMSQQFGNSGLRNLREFIGPDQ SPMLGSPSGNLKSPGTPSGLAGAAAAFYHLKSRSDJAPSSDFDWTSS GFPPTLSSGNSGNLKSPGTPSGTFTSLNTAAPPYQRG LGRRPLLISVAGSQGPASVIPICSSSTFYLNTPEDFT FEPPLQSFGIPPNHKARLTMASPAMLGNVESG GPPPTTASOPASVIPTGSLFTYTMPEDFT FEPPLQSFGIPPNHKARLTMASPAMLGNVESG GPPPTTASOPASVIPTGSLFTYTMPEDFT	1 02.	} =	1	1		1	ASGVSPTLWRPQAAATGLEMPSSGRALLDSP
VGGSRYLSQALSCPHTRLGKLAVVVASY RRPGALAAVPSPLELOBNYDNEYFPERS SQAFRYVLHYYRTGRIHVMEGLCALSFLQEI QYWGIDELSIDSCCRDRYFRRKELSETLDFKK DIEDQESQHESEQDESQCPCFTVRQKLWNIL EKPGSSTAARIFGVISIIFYGVSIINMALMSAEL SWLDLQLIELEYVCISWFTGEFVLRFLCVRD RCRFLRKVPNIIDLLAILFYTITLLVELSGIGOT TQELENVCAHCPGCLRLILAILRAILKAWGR HSTGLRSLGMTTTQCYEEVGLLLIFLSVGISIF STVEYFARGSPEDTTFTSVAWWATTSMT TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI AIINDRFSACYFILKIKEAAVVRQREALKALTK NIATDSYISVFILKIKEAAVVRQREALKALTK NIATDSYISVFILKIKEAAVRQREALKALTK NIATDSYISVFILKIKEAAVRQREALKALTK NIATDSYISVFILKIKEAAVRQREALKALTK NIATDSYISVFILKIKEAAVRQREG STHPASSGPOPWF TASSGFOSMISSNPKVRSSPSGNTGSPKSKQ SGYMVRPFYVMSPSGNPDLDSKFSNGGKQGG ASQSOPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRRSISSADSFDQRDGTPN DDSDIEGENSADHIKSODSQHTPHSMTPSNAT APRSSTPPHGQTIATEPTPAQKTFAKVVYVFS TEMANKAAAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPFRILPQOPAPANO DONSSQNTRLQFTPPTAAVKTPAAVPPRPLDRE SFGVENKLIPSVGSSPASSTPLEPPGGFONSTPN NRAVTPVSQSNSSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGFQQNPGVLDGPQKREGEPT ANAVPTSQSSNSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGFQQNPGVLDGPQKREGEPT ANAVPTSQSSNSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGSGROWPHDAMFGGN AMMAGNSGLIKGFFFFGMPHDHMTTP EQIAVLKLQGFYFEKRKFEQVVVQCSLQ DMMVHQGIGPRGVVRGPPPFYOMFSGGWAG MRLPGFAOMINSEMEGPNVPIPSGRGFF NPCGLSEMFQQQLAEKRKFEQVVVQCSLQ DMMVHQGRGRGVVRGPPPFYOMFSGWAG SWPDDVPKINDGRNSGLRHEFGNPVSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPFEE MKLRFGGSDMLPAGQPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPFEE MKLRFGGSDMLPAGQPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPFEE MKLRFGGSDMLPAGQPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPFEE MKLRFGGSDMLPAGQPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPFEE MKLRFGGSDMLPAGQPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPFEE MKLRFGGSDMLPAGQPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPGRALFRIPQDG YGMGPRFILPMSGOPSNSGLRAILRFIGPDQD YGMGPRFILPMSGAPASVNINGSGNAILFRIPQDPQ RTNSRLSHMPPLPLNESSPTSLANTAPPOVGG LGRKPLDISVAGAGVVSEPGFOPLKSPTMHQVQ SPMLGSPSGNLKSPTVLAKSSELPASPGOWTSS GPPPTASGAASVNINGSLSJSSTYTYMPPEPTL	1	ł	1	1		1	
RRFGALAAVPSPLELCDDANPUNFFFDRS SQAFRYVHYMTGRILAWEQLCALSFLQEI QYWGDEI.SIDSCCZDRYFRRKELSETLDFKK DTEDQESQHESQDFSQGPGYHAKLAWIL EKPOSSTAARIFGVISIIFVGVSIINMALMSAEL SWLDLQLLEILEYVCISWFIGEFVLKELCVRD RCRFLRKVPNIDLLAILFFLVYRD RCRFLRKVPNIDLLAILFFLVYRD RCRFLRKVPNIDLLAILFFLSGISQT TQELIENVGAHCFGCLRLLRAILRMLKAWGR RCRFLRKVPNIDLLAILFSVGISF STVEYFAEQSIPDTIFTSVPCAWWATTSMT TVGYGDIRPDTTIGKIVAFMCILSGILVLALPI AINDRFSSCYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDDFWF ASTRSSGGDDFWF TASSGFOSMISSNPKVRSSPSGNTOSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTFKALFGPGSMGLK NGAGNGAKGKGKERSISSADSFDQRDPGTPN DDSDLECNSADHIKSQDSQHTPRSMTPSNAT APRSSTPPBGOTTATEPTPAQKTFAKVYVYS TEMANKAAEAVLKGQVETIVSFHIGNISNNK TERSTAPLNTQISALRNDFRYDQPPAPANQ DQNSSQNTRLQPTPPIPAPAKPAAPPRPLDRE SPGVENKLIPSVGSSPASSTPLPPDGTRSTPTN — NRAVTPVSQSNSSSADPFAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRQIQRMLFP DEKEPTGAGSGPQONGVLDGPQKKPEGPI QAMMAQSQLGKGFQFRTDVGAPFGPQGHR RVFSPDEMVYPSSNNSQSDFAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRQIQRMLFP DEKEPTGAGSGPQONGVLDGPQKKPEGPI QAMMAQSQLGKGFQFPTTVGAPFGQGHS MRLPGFAGMINSEMEGPNVPNASRFGLSGV MRLPGFAGMINSEMEGPNVPNASRFGLSGV MRLPGFAGMINSEMEGPNVPNASRFGLSGV MRLPGFAGMINSEMEGPNVPNASRFGLSGV SWPDDVPKEPDGNIFPPQGIFSGFGRGERP NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMNRMTPGSGDMLPAGRGREFPPQVGG LSPSRGDFFKGIPPQMGGRELFGMVPSGM KGDVNLNVMMGSNSQMIPQKMREGREFPPQ PSMESSMPPLDVRISSNFTSINTARFPQPQG LGRKPLDISVAGSQVHSPGINPLKSPTLNRAPPQVG LGRKPLDISVAGSQVHSPGINPLKSPTLNRAPPQVG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPAGNIKKSPTPSQLAGMLAGPAAAA SKSSPYLOGAAASSPYLKKSPSLPSATSFGWTSS GPPPTASQASAVNIPGKARAGAGPEE	1		1	1	1		
SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI QYWGDELSDGCCRDRYFRKELSETLDFKK DTEDQESQHESEQDFSQGPCPTVRQKLWNIL EKFGSSTAARIFGVISIIFVGVSIIIMALMSAEL SWLDLQLLELEYVCISWFTGEFVLRFLCVRD RCRFLRKVPMIDLLAILFYITILLVESLGSGYT TQELENVCAHCPGCLRLIAILWMLKAWGR HSTGLRSLGMTTTQCYEEVGLLLIFLSVGISF STVEYFAEQSIPDTFTSVPCAWWWATTSMT TVGYGDIRPDTTTGKIVAFMCLSGILVLALPI AINDRFSACVFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMILRLKGRER ASTRSSGGDDFWF  828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGFQSMHISSNPKVRSSPSGNTQSSPKSKQ EVMVRPTVMSPSGNPQLDSKFSNGGKQGG ASQSOPSPDSSGGHTFRALFGFGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDFGTPN DDSDEECDSDSSGGHTFRALFGFGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDFGTPN DDSDEECDSADHIKSQDSQHTPBSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVVVFS TEMANKAAEAVLKGQVETIVSFHIGNISNNK TERSTAPLATQISALRNDFKFLPQGPPAPANQ DONSSQNTRLQFTPPPAPAFRAAPRPLDRE SPGVENKLIPSVGSPASTPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSGEGGERFDUGGPRGFGGGGF QAMMAGSQSLGKGFGFRTDVGAPFPPCGGF QAMMAGSQSLGKGFGFRTDVGAPFPPCGGF QAMMAGSGSGGPQNFVLDGPKVQCGSLQ DMMVHQHGPRGVVRGPPPYQMTPSEGWAP GGTEPFSDGNMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGRNVPPNASRFGLSGV SWPDDVFKIPDGRNFPPPQGGIFSGFGRGEFP NPGGLSEEMFQQOLAEKGLGJFPPGGGBF PNPGMSEMFMMPGSQRIMMEGNMPFIPRUPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGBVNLNVNMGSNSWIPQKMREACAGFEE MKLRFGGSDMLPAQGKMVPLFFGEHPQG YGMGPRFLPMSQGPGSNSGLRNTREPIGPDQ RTNSRLSIMPPLPLNESSHPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGMPLLSSTPMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAASSVHLKSPSLPASPGWTSS PEPPLQSFGPPPNIKAPLTMASSPALINNESSG GPPPTASQASASVNIHGRSLPSSTFYTMPPEPTL	1	1	1		1	1	
QÝWGIDELSIDSCCADTYRRKELSETLDFKK DTEODESOJESEODPSQCPCTYRQKLWNIL EKPGSSTAARIFGVISIIFVGVSIINMALMSAEL SWLDLQLLEILEYVCISWFTGEFVIRLCVRD RCRFLRKVPNIIDLLAILPFYTILLVESLSGSQT TQELJENVGAHCFGCLRLLRALIKMLAWGR HSTOLRSLGMTITQCYEBYGLLLIFLSVGISIF STVEYFAEQSIPDTTFTSVPCAWWATTSMT TVGYQIRPDTTTGKIVAFMCILSGILVLALPI AIINDRFSSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRESSGODDFW  828 2178 A 6786 5672 1360 GTHPASSGPYPLPAAVSAATREELGEPVPFV TASSGROMEISSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPOLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALFGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDLECCNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTFAKVVYVTS TEMANKAABAVLKGQVETIVSFHINDISNINK TERSTAPLATOJSALRNDFRPLPQOPPAPANQ DQNSSQNTRLQPTPIPAPAKPAAPPRLDRE SPGVENKLIPSVGSPASSTPLPPDGIFNSTPN NRAVTPVSQGSNSSSADPFAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRD(RMLFP) DEKEPTGAGSGPQONFVLDGPQKKPEGPI QAMMAQSQSLGKGPGPTDVGAPFOYGHAR DVPFSPDEMVPPSMNSOSGTIGPPBLDEIMT EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQRGPRGVVRGPPFPYGMAPEGRI PSPRGDEMVPPSMNSOSGTIGPPBLIDIMT EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQRGPRGVVRGPPPFYGMAPEGRI PSPRGDSDMMPHSLPPPGGGIFSGPGRGEFP NPGGLSEEMFQQQLAEKQLGLPPRFEGWAP NPGGLSEEMFQQQLAEKQLGLPPRFEGWAP NPGGLSEEMFQQQLAEKQLGLPRFEGWPGG LSPSRGDFFKGIPPQMGGREREF NPGGLSEEMFQQQLAEKQLGLPRFEGWPGG LSPSRGDFFKGIPPQMGGREREFP NPGGLSEEMFQQQLAEKQLGLPRFEFGWPGG RTMSRLSFMPPLLNRSSNJPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGBNPLKSPTMHQVQ SPMLGSPAGNLKSSPJPTSQLAGMLAGPAAAA SKSSPYLOGSAAASVPHLKSPSLPASPGWTSS FEFPLQSFGPPRIKAPLTMASSAPALONVESG GPPPTASQASVNINGSLJSSSTYTYMPPEPTL	1	1	1	1	ì	1 .	· · · · · · · · · · · · · · · · · · ·
DTEDQESQHESEQDFSQQPCPTVRQKLWNIL  EKPGSTAARIFQVISIIFVQSUBMALMSAEL  SWLDLQLLEILEYVCISWFIGEFVLRFLCVRD  RCRFLRKVPNIDLLAILPFYITILVESLGSIGY  TQELIENVGAHCPGCLRLLRALRAILKAWGR  HSTGLRSLGMTTTQCYEEVGLLLIFLSVGISIF  STVEYFABQSIPDTITGSVPCAWWAATTSMT  TVGYQDIRPDTTTGKIVAFMCILSGILVLALPI  AIINDRFSACYFTLKLKEAVYRQRALKKLTK  NATDSVISVINLRDVYARSMEMLRLKGRER  ASTRSSGGDDEWF  2178  A 6786 5672 1360 GTHPASSQPVPLPPAAVSAATREELGEPVPFV  TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ  EVMVRPTYMSPSGNPQLDSKFSNQGKQGG  SAGSQPSPCDSKSGGHTPKALPGFGGSMGLK  NGAGNGAKGKCKRESISADSFQRDPGTPN  DDSDLKCENSADHIKSOD9HTPHSMTPSNAT  APRSTPPHGGTTATEPTPAQKTPAVVYVFS  TEMANKAAEAVLKGQVETIVSFHIQNISNNK  TERSTAPLNTQISALRDPKPLPQQFPAPANQ  DQNSSQNTRLQPTPPIPAPAPKPAPPRPLDRE  SFQVENKLIPSVGSPASSTPLPPDGTGPNSTPN  NRAVTPVSQGSNSSADPKAPPPPVSGEPPT  LGENPPGLSQCLGLETERSLOTLRDIGRMLFP  DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI  QAMMAQSOSLGKGFQFRTDVAFFCYGGT  QAMMAQSOSLGKGFQFRTDVAFFCYQCLSCLQ  DDMWHQHOPRGVVRGPPPYQMTPSEGWAP  GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ  MRLPGFAGMINSEMGEPNYPSGM  GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ  MRLPGFAGMINSEMGFPYQQLAEKQLGLPFGMAMEGIR  PSMEMINEMFGSGNHEPGONTPIFFRYEGP  LSPSRGDFPKGIPPQMGPGRSLEFGMYPSGM  KGDVNLNVNMGSNSQMIPQKMREAGAGPEE  MKLRPGGSDMLLPAQQKMYPLFFGEHPQQE  YGMGPRFFLPMSQQFGSNSISLNILREPIGPDQ  RTSNRLSHMPPLLPLAHPSSNFTSLNTAPPYQRG  LGRKPLDISVAGSQVHSEGINPLKSPTMHQVQ  SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA  SKSSPVLGSAASPVLKSPSLPAFSPGWTSS  PEPPLQSSGPPNHIKARLTMASPMLGNVESG  GPPPTTASQPASVHLKSPSLPAFSPGWTSS  PEPPLQSSGPPNHIKARLTMASPMLGNVESG  GPPPTTASQPASVHLKSPSLPAFSPGWTSS  PEPPLQSSGPPNHIKARLTMASPAMLGNVESG  GPPPTTASQPASVHLKSPSLPAFSPGWTSS  PEPPLQSSGPPNHIKARLTMASPAMLGNVESG  GPPPTTASQPASSYPLKSSSEPSTYTTMPEYTL			I		1	1	
EKPGSSTAARIFGVISIIFVGGVSIINMALMSAEL SWLDLQLELLEYVSTPGEFVLRFLCVRD RCRFLRKVPNIIDLLAILPFYITLLVESLSGISQT TQELIENVGAHCPGCT.RLIRALIKALKAWGR HSTGLRSLGMTTQCYEVGLLLLFLYSVGISIF STVEYFABQSIPDTIFTSVPCAWWATTSMT TVGYGDIRPDTTTGKVFAMCLS.GGILVLALPI AINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYTSVNILRDVYARSIMEMIRLKGRER ASTRSSGGDDFWF ASTRSSGGDDFWF TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQFSCDSKSGGHTPKALPGFGGSMGLK NGAGNGAKGKGKRESIADSFDQRDFGTPN DDSDIKECNSADHIKSQDSGHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAABAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQOPPAPANQ DQNSSGNTLQVTPTPPPAPARPAAPPPPVSGPPT LGENPDGLSGGQLEHERSIQTLRDIQRMLFP DEKEFTGAQSGGPQONPGVLDGPKKPEGFI QAMMAQSSLCKGPRTDVGAPFGPGGHR DVPFSPDEMVPSMNSQSGTIGPDHLDHMTP EQLAWLKLQQEYTEKRREQTVADGPGKPEGFI QAMMAQSSLCKGPPRTDVGAPFGPGGHR DVPFSPDEMVPSMNSQSGTIGPDHLDHMTP EQLAWLKLQQEYTEKRREPQVVVQQCSLQ DMMVHQHGPRGVVRGPPPYQMTPSEGWAP GGTEPPSDGINMPHSLPPGGMAPHPNMGSQ MRLPGFAGMINSEMEGPNVPNPASRFGLSGV SWPDDVPKIPDGRNFFPQQLEKGPGFTDVGAPFGGGFF NPQGLSEEMFQQLAEKQLGLPPGMAMEG SWPDDVPKIPDGRNFFPQGGFSCPGGREFP NPQGLSEEMFQQCLAEKQLGLPPGMAMEG KGDVNINVNMGSNSMIPGKMEAGAGPE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRFTLPMSQCPGSNSLIRNLEFPIGPDQ RTNSRLSHMPPLINFSSNTTSINTAPPVGG LGRKPLDISVAGSQVHSSGINFLSFTHMIQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGAPAAA SKSPPVLGSAASPVHLKSPSLPAPSPGWTSS PEPPLQSSGPIPNHKARLINASPALGNVESG GPPPTTASQPASVHLKSPSLPAPSPGWTSS PEPPLQSSGPIPNHKARLINASPALGNVESG			1	1	1	1	
SWLDI.QLI.EILEYVCISWPTGEFVLRFIC.VRD RCRFLRÉVPNIIDLLAILPTYTITLLVESLSGISQT TQELIENVGAHCPGCT.RLIRALRAH.KAWGR HSTGLRSLGMTTTQCYEEVGLLLFLSVGISIF STVEYFARQSIPDITIFSVPCAWWAATTSMT TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI AINDRFSACYFTLKLKEAAVRQREALKKLTK. NIATDSYISVNILRDVYARSIMEMLRIK.GRER ASTRSSGGDDEWF  828 2178 A 6786 5672 1360 GTHPASSGPVELPPAAVSAATREELGEPVPFV TASSGFQSMHSSNPKVRSSPSGNTQSSPKSK, EVMVRPPTVMSPSGPDLDSKFSNGGKQGG SASQSQPSPCDSKSGGHTPHSALPGPGGSMGLK. NGAGNGAKGKGKERSISAADSQDRDFTFN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGGTTATEPTPAKTPANVPTSNAT APRSSTPPHGGTTATEPTPAKTPANVPTSNAT APRSSTPPHGGTTATEPTPAKTPANVPTSNAT APRSSTPPHGGTTATEPTPAKTPANVPRPLDRE SFGVENKLIESVGSPASSTPLPPGGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSGGPFP 1 CENPPGLSQEGLERERSLOTLADIGRMLFP DEKEPTGAQSGGPQQNPGVLDGPGKKPEGFH QAMMAQSQSLGKGFGPRTDVGAPFGPQGHR DVFSSPDEMVPPSMNSQSGTIGPDHLDHMTP EQLAWLKLQQEFYEEKRRSLOTLADIGRMLFP DEKEPTGAQSGGPQQNPGVLDGPKKPEGGH QAMMAQSQSLGKGFGPRTDVGAPFGPQGHR DVFSSPDEMVPPSMNSQSGTIGPDHLDHMTP EQLAWLKLQQEFYEEKRRFEQVTVQQCSLQ DMMVHQHFRGVVRGFPPFYVGHPFGWAPEGGWP GGTEPSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKPDGRNFFPQGGFSGFGREFF NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMNMIPGSGGHMEPGNNFIFFRYVEGP LSPSRGDPFKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMHEPGNNFIFFRYVEGP LSPSRGDPFKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMHEPGNNFIFFRYVEGP LSPSRGDPFKGIPPQMGPRELEFGMVPSGM KGDVNLNVNMGSNSQMHEPGNNFIFFRYVEGP LSPSRGDPFKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMHEPGNNFIFFNYVGG PSSRGSDMLKSPQPPSULAGMLAGPAAAA SKSSPVLGSAASPVILKSPSLPAPSPGWTSS PEPPLQSSGNIKSPGLAGMLAGPAAAA SKSSPVLGSAASPVILKSPSLPAPSPGWTSS PEPPLQSSGNIPNHKARLTMASPAMLGNVESG GPPPTTASQPASVHLKSPSLPAPSPGWTSS PEPPLQSSGNIPNHKARLTMASPAMLGNVESG GPPPTTASQPASVHLKSPSLPAPSPGWTSS PEPPLQSSGNIPNHKARLTMASPAMLGNVESG			1	1	1	1	
RCRFLRÄVPNIDLLALIPYTILLVESLSGISGIT TOELVENVGAHCPGCLRLIRALIRMIKAWGR HSTGLRSLGMTITQCYEEVGLLLLFLSVGISIF STVEYFAEQSIPDTIFTISVPCAWWATISMT TUTYGIPDITTIGKIVARMCILSGILVLALPI AINDRFSACVFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDFWF  828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGPQSMHSSPSSNPVKSSPSGNTOSSPKSKQ EVMVRPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGPGSMGLK NGAGNGAKGKGKRERSISADSFDQRDFGTPN DDSDIKECNSADHIRSQDSSHTPHSMTPSNAT APRSSTPPHQGTTATEPTPAQKTPAKVYVYFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKLPQQFPAPANQ DQNSSQNTRLQFTPPPAAFKPLAPAPRPLDRE SPGVENKLIPSVGSSASDPKLPQGPFAPANQ DQNSSQNTRLQFTPPPAAFKPLAPAPRPLDRE SPGVENKLIPSVGSSASDPKAPPPPPVSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGPQQNPGVLDGPCKKPEGIP QAMMAQSQLGKGFGFRTDVGAFFGPQGHR DVPFSPDEMVPFSNNSOSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRKPEQVVVQQCSLQ DMMVHQHGFRGKGPPPPYQMTPSGWAP GGTEPPSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEGPNYPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGFP NPQGLSEMFPQQLAEKQLGLPPGMAMEGIR PSMEMNRMPGSQRIMMERONPIPFRIPVEGP LSPSRGDFFKGIPPQMGPGELEFGMYPSGM KGDVNLNVMMGSNGMPQKMREAAGAGFPE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLLPNSSNFTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSFTMHQVQ SPMLGSPSGNLKSAPVHLKSPSLPAPSPGWTSS PEFPLQSPGIPPNIKAPLTMSSFAMLGNVESG GPPPPTASQPASVNIPGSLSFSSLPAPSPGWTSS PEFPLQSPGIPPNIKAPLTMSSPAMLGNVESG	j	į.	1	1	1	1	
TOELIENVGAHCPGCLRLIKAJIKALKAWGR HSTGLRSLGMTITQCYEEVGLLLIF.SVOISEF STVEYFARQSIPDITTISVPCAWWATTSMT TVGYGDIRPDITTIGKIVARMCLISGILVLALPI AINDRPSACYFILKIKRAAVRQREALKKLITK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDFWF   828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGFQSMHISSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKERSISSADSEDQRDFGTPN DDSDLKECNSADHINSQDSQHTPHSALPGPGGSMGLK NGAGNGAKGKGKERSISSADSEDQRDFGTPN DDSDLKECNSADHINSQDSQHTPHSALPGPGGSMGLK TERSTAPLNTQISALRNDFKLPQGPAFANO DQNSSQNTRLQFTPPIPAPAKPAAPPPPLDRE SPGVENKLIPSVGGSPSSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSADPKAPPPPVSGGPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGGPQNPGVLDGPQKKPEGFI QAMMAQSQSLGKGFGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTGPDHLDHMTP EQIAWLKLQQEFYEKRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPVQMTSEGWAP GGTEPFSDGINMPHSLEPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVFKIPDGRNFFPRQGIFSGFGREFFP NPQGLSEMFGQOLAEKQLGLPPRMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFFRIPVEGP LSPSRGDFFKGIPPQMGFGEFGGNVPSGM KGDVNLNVMGSNSQMPQKMREAGAGFEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPNSSNFTSLNTAPPVQRG LGRKPLDISVAGSQVHSSEGNPLKSFTMHQVQ SPMLGSPSGNLKSFQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEFPLQSPGIPPNIKAPLTMASFAMLGNVESG GPPPPTASOPASVNIFGSLPSSLFSSLPAPSPGWTSS PEFPLQSPGIPPNIKAPLTMASFAMLGNVESG	!		1		i		
##STGLRSLGMTITQCYEEVGLLLLFLSV0ISIF STVEYFAEQSIPDTTTSVPCAWWATTSMT TVGYGDIRPDTTTGKIVAFMCLLSGILVLALPI AIMDRFSACVFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVARSIMEMLRLKGRER ASTRSSGGDDFWF  ###STGLRSLGMTYPYARSIMEMLRLKGRER ASTRSSGGDDFWF  #################################		1				-	
STVEYFAEQSIPDTTFTSVPCAWWWATTSMT TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI AIINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYTISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDDFWF  828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGFOSMHSSNPKVRSSPSGNTQSSPKSKQ EVMYRPFTVMSPSGNPQLDSKFSNQGQGS ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDFGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHQGTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHQNISNIK TERSTAPLNTQISALRNDPKFLPQQPPAPANQ DQNSSQNTRLQPTTPPEPAPKFAAPPRPLDRE SPGYENKLIPSVGSPASSTELPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSNNSQSGTIGPDHLDHMTP EQIAWLKLQGEYFEEKRRKPEQVVVQCSLQ DMWVHQHGPRGVVRGPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKPDGNFFPGQGIFSGFGRGERFP NPQGLSEEMFGQQLAEKQLGLPFGMAMEGIR PSMEMNRMIPGSQRHMEFGNNPIFPRPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEIFPQDE YGMGPRPFLPLNPSSNPTSLNTAPPVQG LGRKPLDISVAGSQVHSPGRINLREPIGPDQ YGMGPRPFLPLNPSSNPTSLNTAPPVQG LGRKPLDISVAGSQVHSPGRINLSPTTMIPQVQ SPMLGSPSGNLSPQTPSGLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKARLTIMASPAMLGNVESG GPPPPTASOPASVNIPGSLPSSTFYTMPPEPTL	i	ł		1			
TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI AIINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMIRLKGRER ASTRSSGGDFWF GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPFTVMSPSGNPQLDSKFSNQGKQGGS ASQSPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDQSHTPHSMTPSNAT APRSSTPPHGQTTATPEPTAQKTPAKVAVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKIPLQQPFAPANQ DONSSQNTRLQFTPPIPAPAFKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSADSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSADFKAPPPPPVSSGEPPT LGENPPGLSQEQLEHERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPSMNSQSTTGPDHLDHMTP EQIAWLKLQQEFYEEKRRFEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGRNVPNPASRPGLSGV SWPDDVPKEPDGNFPPQQGIFSGPGRGEFF NPQGLSEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMPGNNPIFFRPVEGP LSPSRGPFKGGPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRFLFIPMSQQPGSNSGLRNLREPIGPPQ RTNSRLSHMPPLPLNPSSNFTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGIRNLREPIGPPDQ SPMLGSPSGNLLSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKARLTIMASPAALGNVESG GPPPPTASOQPASVNIPGGSLPSSTFYTMPPEFTL	1	1	1	Į.			HSIGLKSLGMIIIQCYEEVGLLLLILSVOISIF
828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV  828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV  TASSGGDDFWF  828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV  TASSGPOSMHSSNPKVRSSPSGNTQSSFKSKQ EVMVRPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALFGPGGSMGLK NGAGNGAKGKGRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHQGTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPPAPAFKPAPPPPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADFKAPPPPPVSSGEPPT LGENPDGLSQEQLEHERSLQTLRDIQRMLFP DEKEPTGAQSGGPQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPGGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQLAWLKLQCEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKTPDGRNFPPQQGIFSGFGRGERFP NPQGLSEMFQQLAEKQLGLPFGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFFKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQCMVFVPLPFGEHPQQG YGMGPRFFLPMSQQPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVMSPGINLREPIGPDQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASFVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKKAPLTMASPAMLGNVESG GPPPPTASSOPASVNIPGGSLPSSTFYTMPPFFTL	1	1	1	1	1		SIVEYFAEQSIPDITFISVPCAWWWAITSMI
828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGFQSMHISSNPKVRSSPSGNTQSSPKSKQ EVMVRPTVMSPSGNPQLDSKFSNQKQGG ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGERSISADSFDQRDFGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPPAPAPKPAPPPPLDRE SPGVENKLIPSVGSSPASTPLPPDGTGPNSTPN NRAVTPVSQGSNSSADPKAPPPPVSGGPPT LGENPDGLSQQLEHRERSLQTLRDIQRMLFP DEKEFTGQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQLGKGPGPRTDVGAPFGPGGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAVLKLQQEFYEEKRRKPEQVVVQCSLQ DMMYHQHGPRGVVRGPPPPYQMTPSGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMGFNVPNPASRPGLSGV SWPDDVPKIPDGRNFPFQGGIFSGPGGEFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMTGSQRIMEPGNNPIPFIPTVEGP LSPSRGDFPRGPQMGFRGEFGFMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MIKLRPGGSDMLPAQQKMYPLPFGEHPQQG TNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVUSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGFAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSFGIPPNHKAPLTMASPAMLGNVESG GPPPPTASOPASNSLPSSTPYTMPPEFTL	1	1	1				
ASTRSSGGDFWF  828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVPFV TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGFGGSMGLK NGAGNGAKGKGKGRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTAQKTPAKVYVYFS TEMANKAAAAVLKGQVETIVSFHQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQQLEHTERSLQTLRDIQRMLFP DEKEPTGAQSGPQQNPGVLDGPQKKPEPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGH DVPFSPDEMVPPSNNSQSGTIGPDHLDHMTP EQIAVLKLQCSFTEKRRPEQVVVQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMGGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRFFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNTTSLNTAPPVQRG LGRKPLDISVAGSAVSPVILKSFSLPASPGWTSS PEPPLQSPGIPPNIKAPLTMASPAMLGNVESG GPPPTASOPASVNIPGSLPSSTPYTMPPEFTL	j	ì		1	<b>!</b>	1	
2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELGEPVFV TASSGFQSMHSSNFKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSPSCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAFKPAAPFRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVITPVSQGSNSSADPKAPPPPVSSGPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGFGFRIDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPPHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNFASRFGLSGV SWPDDVPKIPDGRNFPPQGIFSGPRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPHFRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVINGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLFFGEHPQQE YGMGPPFFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNFTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGFAAAA SIKSPVLGSAAASPVHLKSPSLPAFSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLFSSTPYTMPPSPTL	ì			ļ	ļ		•
TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPFTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGFGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDLKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPIGQTTATEPTPAQKTPAKVVVVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQGPPAPANQ DQNSSQNTRLQPTPPIPAPAKPAPPPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGFNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSGQQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQLAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHFNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQIFSGPRGERFP NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMIRMPGSQRHMEPGNNPIPFRPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGFRPTLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTIMHQVQ SFMLGSPSGNLKSPQTPSQLAGMLAGFAAAA SIKSPVLGSAAASPVHLKSPSLPAFSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLFSSTPYTMPPEPTL			1			ŀ	ASTRSSGGDDFWF
TASSGFQSMHSSNFVRSSPSGNTQSSPKSKQ EVMVRPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGHTPKALPGPGGSMGLK NGAGNGAKGKGKERSISADSFDQRDFGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQPPPAPANQ DQNSSQNTRLQPTPPPAPAFKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVITPVSQGSNSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPHLDHMTP EQIAWLKLQEFYEEKRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPYYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNASRFGLSGV SWPDDVPKIPDGRNFPPQGIFSGPRGERFP NPQGLSEEMFQQQLAEKQLEPGMAMEGIR PSMEMNRMPGSQRHMEPGNNPIFPRLPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVINMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLFFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNFTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SFMLGSPSGNLKSPQTPSQLAGMLAGFAAAA SIKSPPVLGSAAASPVHLKSPSLPAFSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLFSSTPYTMPPSFTL	828	2178	A	6786	5672	1360	
ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGERSISADSFDQRDFTPN DDSDIKEONSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHQNISINKK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAFKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRFGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNPIPFRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MIKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGNPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAFSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASOPASVNIPGSLPSSTPYTMPPEPTL	020						TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ
NGAGNGAKGKGRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVYVYFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQCCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDPFKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MIKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPLPLNPSSNTTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSFTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASFVHLKSPSLAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG	Į.		ì	1		1	EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS
DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQQPPAFANQ DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGFNSTFN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRFFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSSGINPLKSSTIMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASFVHILKSPSLPAPSFGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG			1		1	1	ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK
DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQQPPAFANQ DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGFNSTFN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRFFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSSGINPLKSSTIMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASFVHILKSPSLPAPSFGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG	1		1	1	ŀ	1	NGAGNGAKGKGKRERSISADSFDQRDPGTPN
APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPARKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPPSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPFQGGIFSGFGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGNPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASOPASVNIPGSLSPSSTPYTMPEPTL	1				1		
TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDFKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAPKPAPPPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQLAWLKLQQEFYEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSGWAPP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRFPLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASFVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	1	İ		1	İ		
TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPARKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPPGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQQGIFSGPGRGFFP NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMNRMIPGSQRIMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLAGQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNFTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPVLIGSAAASPVHLKSPLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	}	i	1	1	1		
DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHERESLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDDVPKIPDGRNFPPQGGIFSGPGREERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPVLGSAAASPVHLKSPSLPASPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSSLPSSTPYTMPEPTL		1			1		
SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKKPDGRNFPPGQGIFSGPGGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGFGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MILKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTTMHQVQ SPMLGSPSGNILKSPQTPSQLAGMLAGPAAAA SIKSPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1	i	1	<b>\</b>			DONSSONTRI OPTPPIPAPAPKPAAPPRPLDRE
- NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGNPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1	1 .	1	1	1		
LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPPGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPGGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MIKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSFTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL		'		1	1		
DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	1					-	I CEMBUCI SUEVI EMBED SI ULI BULUDINI ED
QAMMAQSQLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINIMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGREFP NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MIKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREFIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGGSLPSSTPYTMPPEPTL	1	1					
DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGGSLPSSTPYTMPPEPTL	1	1	ĺ		1		DEVEL I QUÁRACIA CA CECEDALIMICA DECEDUCADO
EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPQGIFSGPGRGERFP NPPGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL		1	1	1			
DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPRGERFP NPPQGLSEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPVFSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	1		İ	!	1		DALLOLDENIA LLONGA DE LA CONTROLO CONTR
GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	ļ		1		1	1	EQIAWLKLQQEFYEEKKKKPEQYYYQQCSLQ
MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPFGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	}	}	l	ł	l		
SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1		١,	1	1	1	
NPQGLSEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	l	İ	1	ì	1		
PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL			1	j	. [	1	
PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1	1	1		1		
KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1	1	1		1		
KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL		1			1		LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM
MLKLRPGGSDMLPAQQKMYPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL	1			1	ì	1	
YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPGSLPSSTPYTMPPEPTL			1		1		
RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1				i		
LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	l	1	1	1	1	{	
SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1		1	}	1		
SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL			[	-	ı		
PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1		1		I		
GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1	1	1		1		T
GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL	1	ì			1		1
I DONO DIAMANDIANE DIAMANDIANE DIAMANDIANE		İ	Ì		1		GPPPPT ASQPASYNIPOISLPSSTPYTMPPEPTL
SQINFLSIMIMISATIMINISATIMINISTATIMINISTATIMINISTATIMINISTAT	1					_	SQNPLSIM\MSR\MSKFAM\PS\SNPGYNHDAI

- ARA 944 1	0000	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID		ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod			location	F-Phenylalanine, G-Glycine, H-Histidine,
nuc <b>i</b> -	peptide		in	nucleotide		I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	1	09/496	correspondi	to last amino	O=Glutamine, R=Arginine, S=Serine,
uence	ļ	]	914	ng to first	acid residue	Q=Glutamine, K=Arginine, 5=Serule,
}	ļ	ŀ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ	1	peptide	_	/-possible nucleotide deletion, \-possible
ł	1	1		sequence		nucleotide insertion
ļ		<del> </del>	<del> </del>	Bodanno		KTVASSDDDSPPARSPNLPSMNNMPGMGINT
	l	١	1	·	<u> </u>	QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ
ł	1	ł	1			MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH
1	ļ	l			ł	GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP
·	ļ	1		ļ	1	HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQ
1	1	į .	1			HNGPSGGQGSTPGGWGTTGEGTEGRESIAELQ
i	j .		1.	1	i	SSADAALCKPGGPGGPDSFTVLGNSMPSVFT
	Į.	1	ļ			DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ
ì	ì	1	1	Ī	1	YFPRGEVPGRKQPQGPGPGFSHMQGMMGEQ
i	ļ	ì	}			APRMGLALPGMGGPGPVGTPDIPLGTAPSMP
Į.	i	Į.	ľ		i	GHNPMRPPAFLQQGMMGPHHRMMSPAQST
1	1	1	i i			MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
1		1		1		HPGPVGSPGMMMSMQGMMGP\NRTS
1 205	10150	+	6707	422	3	ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH
829	2179	Α	6797	433	1 3	GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP
1		1				ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG
	1	1		i	1	ASTCPPSPGGSGADKPGFSFFFFSKEARITAG
	1	Į.	1	1	1	AAASSTSSGASCPPVPASSRWGVRSRTRSGSG
1		1				GEREPRDRPSERPRLV
830	2180	Α	6800	3	1911	LPERAFGPRTPRAPRRRRRRLLLSPPPRPPPPL
					1	DREPRAPGPWLCPSRAGTAQDPARIRERRGR
1		1	1	İ		VAGGAAGPAMELRARGWWLLCAAAALVAC
	1		1			ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ
i	1		ļ	1		AEISGEHLRICPQGYTCCTSEMEENLANRSHA
1	ĺ	1		ļ	į	ELETALRDSSRVLQAMLATQLRSFDDHFQHL
	1	1	1	ì		LNDSERTLQATFPGAFGELYTQNARAFRDLY
	ł	1	-	}	ļ	SELRLYYRGANLHLEETLAEFWARLLERLFK
	1	1	1	1	i	SELKLY I KOANLINEET LALI WARDERON
	1	1	1 -	ĺ		QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\
1		1	l l	1	1	RELRLRAT\RA\FVAAR\SFVQGLGVAS\DVVR
ĺ	í		1	i		KVAQVPLG\PEC\SRAVIEAGSYC/ALHCVGVP
1	1	1	1	1	1	GARPCPDYCRNVLKGCLANQADLDAEWRNL
ì		i i	1		1	LDSMVLITDKFWGTSGVESVIGSVHTWLAEA
	1	1	ı		1	INALQDNRDTLTAKVIQGCGNPKVNPQGPGP
1	1	1	i			EEKRRRGKLAPRERPPSGTLEKLVSEAKAQL
Į.	1	1			1	RDVQDFWISLPGTLCSEKMALSTASDDRCWN
í	ſ	ĺ				GMARGRYLPEVMGDGLANQINNPEVEVDIT
	ŀ		1	l	1	KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF
į.	1		ļ	1		QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS
İ	1	i		ļ		ODASDOGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
		1				SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL
	1	1		1		LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA
1	1					EGRLREKLFSGYDSSVRPAREVGDRVRVSVG
ì				1	Į.	LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS
1			İ	Į	1	WDPAEHDGIDSLRITAESVWLPDVVLLNNND
ļ	1	1		1	1	GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS
1				1	1	IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT
			ì	Į.	1	GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS
1		- 1			1	RLIQPPGDPRGGREGQRQEVIFYLIRRKPLFY
	1	-	1		1	LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF
	1	1		1	1	LVINVIAPCILITELATE VETETEL CUDITIVVI MET
	1			1		ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT
1	1	- 1		1	1	MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV
1		- 1		1	1	RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP
	1		1	1		GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL
		Ì	1	ſ		SAPDLRRFIDGPNRAVALLPELREVVSSISYIA
1		}	}	1		RQLQEQEDHDALKEDWQFVAMVVDRLFLW
	1	- 1	1	1	ļ	TFIIFTSVGTL\VIFLDATYHLPPPDPFP
L					1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK
832	2182	A	6824	71	1079	ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY
	1	1			1	DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN
·	1	-		1	i	DMEMITIONUEKAKI IMEMPYIKI ELIKOUN
}	1	}	İ	ì	1	GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI
		1				KVIPEFSEPEEEIDENEEITITFFEQSVIWVPAE
1	1	1		- 1		KPIENRDFLKNSKILEICDNVTMYW\INPTL\IS
L						

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide eletion, \=possible nucleotide insertion  GTFAKQLHHNFAPIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV
833	2183	A	6846	116	602	IMPCNWWYARMLGRV  EAEGEQVCGAKCCGDAPHVENREEETARIGP GVMESKEERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFC\LMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSRESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMI.DCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVV LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6855	315	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGKYRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG
						RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP
837	2187	A	6863	2	1615	VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHINLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGANLNARTSMDEMPIDLCEEEEFKVL LLELKHKHDVIMKSQLRHKSSLSRRTSHRQA S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

					CK Track	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Ammo acid sequence (A=Alamine C=Cysteme,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	ucioc	1	1		acid residue	O=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first		
	ł	ł	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]	ł	1	i .	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	l .	}	i	peptide		/=possible nucleotide deletion, \=possible
1		i		sequence		nucleotide insertion
				Sequence	ļ	PNPRLEK\PVLLSEFPTKIPRGELDMPVENGLR
1	ŀ	4	1	1	İ	
ì	Į	1	ł		ł	APVSAYQYALANGDVWKVHEVPDYSMAYG
1	ł	ł	1		1	NPGVADATPPWSSYKEQSPQTLLELKRQRAA
l	1	1			1	AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI
	İ	1				GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI
	Į	1				EEMEEKVHGCCRIS
<u></u>	ļ <u>.</u>	<u> </u>	<del> </del>			
838	2188	Α	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS
1	1	l				LRPRKLDFFRSEKELNHLAVDEASGVVYLGA
	Ì	[		1		VNALYOLDAKLQLEQQVATGPVLDNKKCTP
l	1		1		j	PIEASOCHEAEMTDNVNQLLLVDPPRKRLVE
1		1	į.		l .	CGOLLKGI\CALRALSNISLRLFYEDGSGEKSF
1	ì	1	1	1	I	1 `
1	1	1				VASNDEGVATVGLVSSTGPGGDRVLFVGKG
		1	1	1		NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT
		1	1	1	1	YKAGYLSTNTQQFVAAFEDGPYVFFVFNQQD
		1		1		KHPARNRTLLARMCREDPNYYSYLEMDLQC
1		l.	1		1	RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF
1		i	1			SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN
1.		İ			1	
1	1	1	1	1	1	ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK
	1 .	1	1	1		SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN
1			1 -			LTAVTVAAENNHTVAFLGTSDGRILKVYLTP
1		1		1		DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY
	1	1	İ		1	AMTODKVFRLPVQECLSYPTCTQCRDSQDPY
1		1	1		1	1
	1	1	1		i	CGWCVVEGRCTRKAECPRAEEASHWLWSRS
İ	1					KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA
	1	1	1			LSEEDELLCLFGESPPHPARVEGEAVICNSPSS
1	1 .				1	IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF
			i	1		YDCRQAMSLEENLPCISCVSNRWTCQWDLR
1		1	1			YHECREASPNPEDGIVRAHMEDSCPQFLGPSP
	1	1		ļ		
	1	ĺ	1			LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD
	1	1	ļ	1	1	LLKFMEPVTMQESGTFAFRTPKLSHDANETL
l.	1	1	1	1	•	PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC
1			İ	1		SLCRAANPDYRCAWCGGQSRCVYEALCNTT
J		1	ļ	1	1	SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ
1		1	1		į	· · · · · · · · · · · · · · · · · · ·
1		1	1	Ţ	ľ	AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA
	1	1	1	1		AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP
1	-	1	1		Ļ	KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED
1	1	1	1	1	1	VRVTLNGVPCKVTKFGAQLQCVTGPQATRG
1		1	1	J		OMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE
	1		ŀ	L		PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP
1		ĺ		T		
1	1	1		1	1	LQSWQPPREAESLQPMTVVGTDYVFHNDTK
}	1	1	1	1	1	VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT
1	ı	l		1	1	EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR
1	1	1			1	GTNLNKAMTLQEAEAFVGAERCTMKTLTET
1	1	i	1	1	1	DLYCEPPEVOPPPKRROKRDTTHNLPEFIVKF
1	1		1		1	GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM
	1		1		1	
1		1		1		VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG
1	1	]	1	1	1	LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG
1		l			1	IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL
Į.	1	1 '			1	DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT
		1		1	1	L\ENOPEFSARAKVYFASLLTVALHGKLEYYT
	1				1	
1	1	1				DIMHTLFLELLEQYVVAKNPKLMLRRSETVV
i	1		ĺ			ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI
1	1					KHQVEKGPVDAVQKKAKYTLNDTGLLGDD
1	ı	1				VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ
	1	ļ	1	[		VKEKIIDOVYRGOPCSCWPRPDSVVLEWRPG
ı		1		1		
1			1	1	1	STAQILSDLDLTSQREGRWKRVNTLMHYNVR
l	1	1	1	1	1	DGATLILSKVGVSQQPEDSQQDLPGERHALL
ı	1	1	ł	· ·	l	EEENRVWHLVRPTDEVDEGKSKRGSVKEKE
E .	1	1	1	1	1	1
1		Į.	1	1	1	RTKAITEIYLTRLLSVKGTLOOFVDNFFOSVL
						RTKAITEIYLTRLLSVKGTLQQFVDNFFQSVL APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HIWKTNSLPLRFWVNILKNPHFIPDVHVHEVV DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL YAKEISTYKKMVEDYYKGIRQMVQVSDQDM
1			-			NTHLAEISRAHTDSLNTLVALHQLYQYTQKY YDEIINALEEDPAAQKMQLAFRLQQIAAALE NKVTDL
839	2189	A	6872	1	1485	RARRLALQCHVCVCALTPGEQSGRRLPGQT WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH EDQTDCSSLRDENNKENYPDAGALVEEHAPP SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP KSIFKAESGRSHGESQETEHVVSSQSECQVRA GTPAHESPQNNAFKCQETVVRLQPRIDQRTAT SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS TQSVLA\DGTDSADPSPVHKDGQNEADSAPE DLHSVGTSRLLL/YHITDGDNPTAVRHGCSL/F SGQSQRFNLDPESAPSPPSTQGFMMPRSSSRC SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLLCTDV
840	2190	A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN YLKQPVVKEKEKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRLKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVHPERESNQEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTOKHLSNREEMELEHSDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII
841	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRGSTSLPGDKKSG TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methioriine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA
						VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS LPSWRARRLSLSRSLVLNRLRPVASGGGKAQ PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAIIRQARQRREKREYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC PDFARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHE\APSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
842	2192		6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSRKTRHNSTDLPL
843	2193	A	6919	2	663	AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195	A	6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVITSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLIPHFKPWLHPEQSP LPSLALSIELSVQHADSILENIDESAVAESREE RIMGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

PCT/US01/03800

NO. of   No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide seq- uence uence  914  914  915  846  2196  A 6944  42  2672  872  872  874  874  875  875  875  875  875  875					l .		
### B47 2197 A 6951 3 1994 ### B47 2197 A 6951 3		peptide		in	nucleotide	location	
### Bin to first smin seids of peptide residue of peptide peptide peptide sequence    ### Bin to first smin seid of peptide sequence    ### Bin to first side of peptide sequence    ### Bin to first sid	eotide	seq-		USSN	location		
mino acid recidud of peptide sequence pitches and the company of t	scq-	uence					
Peptide   Sequence   Y=Tyrosine, X=Unknown, *-Stop codon,	uence	ļ		914			
Poptide   A-possible nuclotide delation, massible   nucleitide insertion   nucleitide   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide insertion   nucleitide   nucleitide insertion		i					
196		}	ļ			sequence	
846	1	į	1	l		Ì	, , , , , , , , , , , , , , , , , , , ,
BLITLGETQEEDBLIPKDYRSLDYDRCINDP YLEVLEYMDNIKGBRYEAVKWMYYAIGV CTGLVGLFVDFYRLFTQLKFGVYGTSVECCS QKGCLALSLELLIGHTVFHCESLIGLBEYE AGSGTFGKCYLYARQVGGLVRIPTLLWKAL GVLLTVAMLLIGLGSPMISGSVVQKAGLPQ POSISLRKJQPHFYPRSDRYGKJDKRDFVSAG AAAGVAAFGAPIGTIFSLEGESSPWOGL TYMKYLFCSMSATYTLMFFRSGIGFGSWGSFOL POLINGEFKCSDSDKKGLIBVTAMOLGFV VMGVYGCHLGATFMCLNKELAKYEMRNYEP REXLYVRISLLVSLYTTVVVFVASMVGGFF ROMSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNGGSALQLFHQDGTFSPV TLALFFYLYTHLACVTYTVVFVASMVGGE AAAGGRIVANVLSSYIGLGHTSGTFALIGAA AFLGGVVGNTMISLFVHJESNEETTGLFMYT LUNGKWTODFFKKGNYDHVGLRCVPLLEW ETEVEMBULRASDIMBFYLTYVPFHLRGSLY SILRTTVHAFPVYTENGGREEMKGNQLIS NNIKFKKSSUTRAGGRGKRSGSNKSYSFSEL RNMCDEHLASEPAEKEDLLQOMLERRYTFY PRLYPDGSSEDVTMERRFRPLTHFULLESNOETTGLLRGS LVTTLVRGVCYSEGSSASQFBLSYAEMAED YRRYDDHOLDLTLNFRMIDVDTFYNNOFFF TVSNTHVSQVFNLFRTMGLRHLPVNAVGE LYTENDHOLDLTLNFRMIDVDTFYNNOFFF TVSNTHVSQVFNLFRTMGLRHPVNAVGE LYTENDHOLDLTLNFRMIDVDTFYNNOFFF TVSNTHVSQVFNLFRTMGLRHPVNAVGE LYTENDHOLDLTLNFRRFRGARDOFFVK VPYRKITINFGCVUDGMPGVVFKAPGYLEI SMRRILBAAFFIKTVRPLPGLEISGEYST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVGNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVGNIST CLICKRSMSVSKEYNLRHPYTHTSELGEFST VGKRKIOGEGRVPQEKWERAYFFVEVGNIST CLICKRSMSVSKEYNLRHPYTHTNICHLARDFT VGKRKIOGEGRVPQEKWERAYFFVEVGNIST CLICKRSMSVSKEYNLRHPYTHTNICHLARDFT VGKRKIOGEGRVPQEKWERAYFFVENGL NETHT TRANSLAGHTER SMRRILBAAFFIRTSELGEFST VGKRKINGRAFT VGFRST VGKRKINGRAFT VGFRST VGKRKINGRAFT VGFRST VGKRKINGRAFT VGFRST VGKRKINGRAFT VGFRST VGFRST VGFRST	846	2196	A	6944		2672	
CTGLVGI_PUPPFVLIFTQLKFGVVQTSVEECE  QRCALSLISLELIGNETVELSSLIGLEEPVE AGSGTTEGKCYL_YARQVFGL_VELPTLLWALL GVLLTVAAMLLIGGTSLTSLEGSSFWNQGL POSISLRXIGPNEPYFESDRYGKDKDDFVSAG AAAQVAAPGGTIFSLEGSSFWNQGL TVKVLIFCSMSATTTLNFFRSGIQFGSWGSFG-OF- POSISLRXIGPNEPYFESDRYGKDKKDLWFVSAG AAAQVAAPGGTIFSLEGSSFWNQGL TVKVLIFCSLSTLLNFFRSGIQFGSWGSFG-OF- POSISLRXIGPNEPYFESDRYGKCMLWADLGFEV VMGVUGGLLGATRNCLNKTLAKYRMRNVIFF KPKLVRVLESILVSLYTTVVVPVSAMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESALQLFHQGGTFSFV TLALFFVLTFILAGWTYGISVPSGLOF GAAFGRLVANVLKSYIGLGHYSGLOF UTLLVGGKWTGDFNRGNVDIHVGLRGVFLLEW ETEVEMBKLRASDIMFENTLYVYPTRIQSLV SILRTTVHHAPPVVTENRGNEKEFMKGNQLIS NNIKFKSSLTRAGEQKRASGSMKSYFSSEL RNMCDEHLASEPFAKEDLLQGMLERRYTY PRLYPDGSPSEDWTMEERFFLTHGLLRSQ LVTLLVRGGVYGSGSSASGPRLSYAEMAGD YFRYPDHEDLLTLLNFRMVDVDTYPMNPSFF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE LVTLLVRGGVYGSGSSASGPRLSYAEMAGD YFRYPDHEDLLTLLNFRMVDVDTYPMNPSFF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE LSIEKIKQLREGVNDLSKRFGEALGVDFPVX VPYRXITTNPGCVVIDGMPFGVVFKAPGYLLE SSMRRILBAASFIKTVURFLGGLISNGEYST VGERKIDGEGRVPGKKRAFTFVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGEYST VGERKIDGEGRVPGKKRAFTFVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGEYST VGERKIDGEGRVPGKWRAPFTVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGEYST VGERKIDGEGRVPGKWRAPFTVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGEYST VGERKIDGEGRVPGKWRAPFTVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGEYST VGERKIDGEGRVPGKWRAPFTVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGEYST VGERKIDGEGRVPGKWRAPFTVEVQNIST CLICKRSMSVSEFVURFTVINFLGGLISNGETHL DSYGSSKORLPQLSSDWRDLAFLVDMTM HNALNISLOGISDVTQMYDLBFALAKLCL WETHT TRNNLAHFFTLKLVGRNESDGLNYPF KLAELKTEPGKLASDFLVAKRFFSLEEL DSYMSSKORLPQLSSDWRDLAFLVDMTM HNALNISLOGISDVTQMYDLBFALAKLCL WETHT TRNNLAHFFTLKLVGRNESDGLNYPF KLAELKTEPGKLASDFLVAKRFFSLEEL DSYMSSKORLPQLSSDWRDLAFLVDMTM HNALNISLOGISDVTQMYDLBFALAKLCL WETHT TRNNLAHFFTLKLVGRNESDGLNYPF KLAELKTEPGKLASDFLVAKRFFSLEEL DSYMSSKORPLPQLSSDWRDLAFLVDMTM HNALNISLOGISDSRUPMATH KLAELTFEFCRLASDFLVAKRFFSLEEL DSYMSSKORPLPQLSSDWRDLAFLYDFROTTVC CRNMTKHPQLSSYTRITNIGCKKEAVM KLAELTFEFCRLASDFLVAKRFFSLEEL DSYMSSKORPLPQATATT HAMPONDATATT HAMPO	***	1 -1,70	1			}	ELTILGETQEEEDEILPRKDYESLDYDRCINDP
OKOCLAISLIELLGFNLTFVELSILGLEPVE AGSGTTGKCTVLARQVPGIVELTLIWKAL GYLLTVAAMILIGLGSPMINSGSVVGAGLD POSISLRIJOPNEPYFSSRYKGKDKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNGGLI TWEVLPCSMSATFTLNEFSSIGGOGSWGSPOL POLINFGERKCSSDEKKCHLWTAMDLOFFV VMGVIGGLI.GATFNCLNKRLAKTRAINNYH REKLVRLESLUSLUTTVVFVFASMVLGGE ROMSSSSQIGNDSPQLQVTEDVNSSKTFFEO- NDTYMDMATLFFNQESALQLFHQDGTFSFV TLALFFVLYFLLACWTYGISVPSGLFVFSLU- GAFAGRUANVLKSYVIGLGHYSASMVLGGE ROMSSSSQIGNDSPQLQVTEDVNSSKTFFEO- NDTYMDMATLFFNGESALQLFHQDGTFSFV TLALFFVLYFLLACWTYGISVPSGLFVFSLU- GAFAGRUANVLKSYVIGLGHYSASMVLGGE ROMSSSSQIGNDSPQLQVTEDVNSSKTFFEO- NDTYMDMATLFFNGESVIGLGHYSTFALGAA ARLGGVVRMTISLTVLLESTNEITYGLFMVT- LAMPGVATTHLAGAA ARLGGVVRMTISLTVLLESTNEITYGLFMVT- LAMPGVATTHLAGAA ARLGGVVRMTISLTVLLESTNEITYGLFMCAPLAGAAA ARLGGVVRMTISLTVLLESTNEITYGLFMCAPLAGAAA ARLGGVVRMTISLTAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1			1	[		YLEVLETMONKKGRRYEAVKWMVVFAIGV
AGSGITEGKCYLYARQÜPGL VRI_PTILLWAK_ GVLLTVAAMLLIGGSMHISSSVYGAGLPQ POSISLRXIGPNEPYPESDRYGKDKDEPVSAG AAAGVAAAPGGITIFSLEESSESWINQGL TWKVLPCSMSATPTLNFFRSGIQFGSWGSFQL POLINFGERKCSDSDKKCELWTAMDLGFPV WMCVIGGLLGATIFCLKRIKARYRINNYEP KRKLVRVLSSLVSLVTIVVVFVASMVLGGE ROMSSSQIGNDSSSQIGNDSSQLQVTEDVNSIKTFFCP NDTYNDMATLFFNPQESALQLFEQDGTTSPY TLALFFVLYFLLAGWTYGISVPSGLEVPSLLC GAAFGRLVANVLKSYIGLGHTYSGTPALIGAG AFLGGVVRMTISLTVLIESTINGETYGLPMVT LMVGKWTGDFFNKGNYDIHVGLRGVPLLEW ETEVEMDKLARSDIMEPNLTVYPTIRGSLV SILRTTVHHAPPVVTENRGNEKEFMKGNQLIS NINKFKSSILTRAGEQKKRSGSMKYSPSEL RNMCDEHLASEPAKEDLLQOMLERRYTYGLPMVT PRILYPDGSPSDDWTMEEFREPLTFIGLLRSQ LVTLLVRGVVCYSSQSSASQPLSYABMAED YPRYPDHDLDLTLLNPRNIDVTPYMPSPF TVSPNTHVSQVPNLFRTMGLRHELVVNAVGE IVGITTRINLTYEFLQARLAGHVQTI  847 2197 A 6951 3 1994 MYNSSSVTNSAAGVBDLNIVQVTYPDMEKER LSIERKIQLREQVNDLFSKRFGEAGVDTPVX VPYKRYDHENDLTLLNPRRMGLRHELVVNAVGE VPKYRDHDLDLTLLNPRRMGLRHELVVNAVGE VPKYRDHDLDLTLLNPRRMGLRHELVVNAVGE VSYRKTENDGGCVVCVDGMPGOVYKAPGVLLE SSMRRILEAAEFIKFTVRHJGLELSNGCYST VGKRKENDGGGRVPGKKWRAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHTVOG MERMREDKLHELKKGLRXYLLGLSDTECPE QKQVFANSFTQKSRVGVPTDLAGNLWELKE EKIESFVAYSIALDETDINNTTOLAGSTURE LEKIESFVAYSIALDETDINNTTOLAGSTURE LEKIESFVAYSIALDETDINNTTOLAGSTURE LEKIESFVAYSIALDETDINNTTOLAGSTURE LEKIESFVAYSIALDETDINNTTOLAGSTURE LEKIESFVAYSIALDETDINNTTOLAGSTURE LEKIESFVAYSIALDETDINNTTOLAGTLAGNLWELKE NECHNISKLUSVASTGTPPAVVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGTPPAVDANNGLVTKL NECHNISKLUSVASTGT LEGERSTATTNOCHNISKLUS NATURE  BAB	1	1	ł	ł		1	
GYLLTVAAMILINGLÖSPMIHSGSVVGAGLIQ POSISILRIJOPHPYPSDRYKGKRDFVSAG AAAGVAAAFGAPIGGTI-FILEGGSSFWNQGL TWKVLPCSMASTTI-NFFISGIPGSWGSPQL POLLINFOEFKCSDBDKKCIL WTAMDLOFFY VMCVIGCLLGATHOLINKTLAKYRARNVIFP KPKLVRVLESILVSLVTTVVVFVASMVLGEG RQMSSSGGNDSPQLQVTEDVNSSIKTFPCP NDTYNDMATI-FIVRPGSALQLFRQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVRSILC GAAFGRLVANVLKSVIGLGHYSDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVRSILC GAAFGRLVANVLKSVIGLGHYSDGTFALIGAA AFLGGVVRMTISLTVLLESTNETTYGE.PMVT- LAVIGKWTGDFFRRGINVDHVGLRGVPLLEW ETEVEMPKLRASDIMEPRILTYVPPHTRICGSL SIRRITVHAHPPVVTSTRORRIKEEMKONQLIS NNIKFKKSSILTRAGEQRKRSGSMKSYPSSEL RPMCDEHASEEPFAKEDLQQMLERRYTFY PPLYPDGSSEDWTMEERFRPLTFHGLILRSQ LVTLLVRGVCYSESGSSAGVPRSJCANAMADD YPRYPDHIDLDLTLLNFRMIJOVPTYMNFSEF TVSPRTHVSQVTNLFRTMGLRHLPVVNAVGE PVRTYPDHIDLDLTLLNFRRWIDVTYPYMNFSFF TVSPRTHVSQVTNLFRTMGLRHLPVVNAVGE PVRTYPDHIDLDLTLLNFRRWIDVTYPMNFSEF TVSPRTHVSQVTNLFRTMGLRHLPVVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVNAVGE VGRTRTMVSQVTNLFRTMGLRHLPVNAVGE VGRTRTMVSQVTNLFRTMGLRHLSVTNLFTMGLLRSQVTNLFTMGLLRSQVTNLFTMGLLRSQVTNLFTMGLLRSQVTNLFTMGLLRSQVTNLFTMGLLRSQVTNLFTMGLLRSQVTNLFTMGLLSGNTLFTMGLTMGLTMGLTMGLTMGLTMGLTMGLTMGLTMGLTMGL	<b>!</b> .	ĺ	1	1		İ	1 \
FOSISLIKIJONPPYRSDRYGKUKERDFVSAĞ AAAQVAAAGPAGIGITINELGSSFWOQGI. TWKVLPCSMSATPTLAFFRSGIGFGGWGSPWOQGI. TWKVLPCSMSATPTLAFFRSGIGFGGWGSFQL PGLIAFGEFKCSDSIKKCHL WTAMDLGFFV VMGVIGGILGATFNCLNKKLAKYRMENVIFP KFKLVRVLESLLVSLVTTVVVFVASMVLGEC RQWSSSQIGNDSFQLQVTEDVNSSKTFFCP NDTYNDMAILFFNPQESALQLFHQDGTFSFV TLALFFVLYPLLACWTYGISPSYGILFYSGLF GAAFGRLVANVLKSYIGLGHYSGTFALIGAA AFLGGVVRMTISLTVLESTNAETTYGLPMVT. LMVGKWTGDFFNKGRYDHYGLRGVFPSLLC GAAFGRLVANVLKSYIGLGHYSGTFALIGAA AFLGGVVRMTISLTVLESTNAETTYGLPMVT. LMVGKWTGDFFNKGRYDHYGLRGVFPSLLC GAAFGRLVANVLKSYIGLGHYSGTFALIGAA AFLGGVVRMTISLTVLESTNAETTYGLPMVT. LMVGKWTGDFFNKGRYDHYGLRGVFPSLLC SILRTTVHHAFPVYTENGREKEFMKGNQLIS NNIKFKKSSLITRAGEQRKRSQSMKSYPSSEL RNMCDEHASEEPAEKEDLLQQMLERRYTPY PRLYPOGPSEDVTMEERRFTTFGILRSQ LVTILLVRGVCYSESQSSASQPRLSYAEMAED YPPLYPDHDLDLLTLNPPMUTPYTMPEPSP TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGHTINLTVFELQARLRQHYQTT VSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGHTINLTVFELQARLRQHYQTT VSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGHTINLTVFELQARLRQHYQTT VSSSVTINSAAGVEDLNIVQTVPFNEKER LSSIEKIKQLREQVYDLFSRKFGGALIGVDFPVK VPYRKITENGEGVPQEKWERAFYEVQNIST CLICKRSMSVSKEYNLRRHYQTNINSKHYQQY MEKRMDEKLHFELKKGLRRYLGLSDTECPE QKQVANPSPTOKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIADETIDNATYGLSFRGVDE NFDVSEBILDTVPMTGTRSGNBIPSRVEKSILK NFCCHWSKL VSVASTGTPPMVANGILVTKL KSRVATTCKGAELKSICCHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNISESTVLRKSLK KSRVATTCKGAELKSICCHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNISESTVLRKSLK KSRVATTCKGAELKSICCHPESLCAQKLKM DHVMDVVKSVNWICSRGLNISESTVLRUSH LBGQYGSLLYTTEKWLSRGLNISESTTILLYEL DSQYGSLLYTTEKWLSRGLNISHERTFTLLYEL DSQYGSLLYTTEKWLSRGLNYDLKALIKFTPLPSKTK LSSPKSTRCHCHORNYDLKALIKFTPLPSKTK DSVHEELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHICAKLIKSMFGSTYICEG LFSIMKLSKTKYCSQLKDSQWDSVLHAT SOWYLEGLOMEDSTYRTNISTDFTLLCFPVGPRIQM LLEGLFSLFVYTHLASTDAPLMKPTPLPSKTK ASAPVQCLLLMAATSSQGLAKPHSGTTPTC CFANATKEPIORLESTYRTNIYDCHLAVINLIF AKPISLAGRAMMQTTLTSTFAVTECLLLVVMS YDLYVALCHLPTKALMATMVCTILAVTSWT TGVLLSLHLVALLLPLFCPCPKYMFTFCELLA VLKLACADTHINEMMVLAAGAIGGLVGPLSTV TGVLLSLHLVALLLPLFCPCPKYMTFTCELLA VLKLACADTHINEMMVLAAGAIGLVGPLSTV		į	ł			1	
AAAGWAAFGAPIGGITISLEEGSISPWOGL TWKVLPCSMSAFTILAFRESIGFGSWGSFQ- PGILNFGEFKCSDSIDKKCHLWTAMDLGFV VMGVIGGLIGATFPICLNKELAFVARMENVER KPKLVRVIESLLVSLVTTVVVFVASMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSKIFFPQ- NDTYNDMAILFFNPQESALQLFHQDGTFSPV TLALEFVLYSLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYGIGUPTSTPALGA- AFLGGVVRMTISLTVILESTNÆTTYGLFNYT- LMVGKWTGDFFNKGNYDHVGLRGVPFLEW ETEVEMDKLRASDIMEPNLTVVPHTRIGSLV SILKTTVHHAFPVTENRGRESFMKGNQLIS- NNIKFKKSSLTRAGEGRKRSQSMKSYPSSEL RNMCDEHASREFAREDLIQMERRYTFY PRLYPDGSPEDWTMERFRPLTFHGLLRSQ LVTILLVRCVCYSESQSSASQFRLSYAEMADD YPRYPDHDLDLTLNPFMUVOUPTYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGITTRHLTYPELQARRAGYOTI SEKENGULESVYNDFSKRSGALGVOFPVK VPYRKITDFQCVVIGGMPGVVVFKAPGYLEI SSMRRIEAAEFKFVTRPLPGLSNGEYST VGRKKDDGGRVPGEKWERAYFFVEVONIST CLICKRSMSVSKEYNLRRHYOTNHSKHYDQV MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQVPGLAGNLWBKLR EKIRSFVAYSJALDETIDNNTJCAJFIRGVDE NFDVSEBLLDTVPMTGTKSGGEIFSRVEKSLK NFCCHWSKLVSVASTGTPPMVDANNGLVTKL SKRVATFCKGAELKSICCHPESLCACKLKM DHVMDVVVKSYNTCSRGLHALVDMTM HLNALNISLQGHSQIVTOMYDLRAELAKLCL WETTITRHLSTALFPTLKLVSHALVDMTM HLNALNISLQGHSQIVTOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSQUTVOMYDLRAELAKLCL WETTITRHASTDALALVDMTM HLNALNISLQCHSLAVNICTSGRIALVOMTHANT HLNALNISLQCHSLAVNICTSGRIALVINGTPTTVC CFNANTKLYFORLESYTRINGTCHSAV  SVOYLFGRFTFTHASTDALALUKPTPLPSKTK ASAPVOCLLLMAATTSPQGLAKPHSGTTTTVC CFNANTKLYFORLESYTRINGTCCFKEAW  TGVLLSHHLVAULLPLPFCPCKYLHTFCELLA VLKLACADTHENEMVALAGAISGLVGPLSTV VLKLACADTHENEMVALAGAIS	1			Ì	l		
TWKYLPCSMSATPTLNFRSGIGFGSWGSFQL POLLNFGERKCSDSNKCHL WTAMDLGFFV VMGVIGGILGATFNCLNKRLAKYRAKNVLP KPKLVRVLESLLVSLVTTVVVFVASMVLGBC RQMSSSQIGNDSFQLQVTEDVNSSIKTFPCP NDTYNDMATLFFNPQESAUQLFHQDGTFSVF TLALFFVLYPLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHTYSGTFALIGAA AFLGGVVRMTISLTVILESTNETTYGLPMYTI LMYGKWTGDFFNKGRYDHYGLRGVGPILEW ETEVEMDKLRASDIMEPNLTVYPHTRIQSLV SILATTYHHAFPVYTENRGNEKEFMKCNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNNCDBHASSEPAEKEDLQMLERRYTFY PPLLYPDQSPSEDVTMERRRPLTTHGLILRSQ LVTLLVRGVCYSESQSSASGPXSYAEMAED YPRYPDHDLDLTLLNPRMIVDVTPYMNFSPF TVSPNTHYSQVFNLFRTMGLRHLPVVNAVGF INGITTRINLTYEFLQARLRQHYQTI A 6951 3 1994 NTNSSSVTISAGAGVEDLNIVQVTYPDDEKER LSSIEKKQLREQVVDLFSRK*GEAIGVDFPVX VPYRXITIFNGCVVDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVRPLPG-ELLSNGEYST VGERKDDGGRVPQEKWERAYFFVEVGNIST CLICKRSMSVSKSYNLRRHYQTNNSKHYDQY MERMRDEKLHELKKGLRSYLLGLSDTECPE QKQVANPSPTOKSVPYVLEGLLSNGEVST VGERKDDGGRVPQEKWERAYFFVEVGNIST CLICKRSMSVSKSYNLRRHYQTNNSKHYDQY MERMRDEKLHELKKGLRSYLLGLSDTECPE QKQVANPSPTOKSVPYVENGLANLWEKLR EKIRSFVAYSIADDETIDNNTTQLAFFRGVDE NFDVSEELLDTVPMTGTKSGGANLWEKLK EKIRSFVAYSIADDETIDNNTTQLAFFRGVDE NFDVSEELLDTVPMTGTKSGGANLWEKLK EKIRSFVAYSIADDETIDNNTTQLAFFRSVEFSLEEI DSSMSRKCKPLQLSDIBWRIGANLWEKLR EKIRSFVAYSIADDETIDNNTTQLAFFRGVDE NFDVSEELLDTVPMTGTKSGGANLWEKLK EKIRSFVAYSIADDETIDNNTTQLAFFRGEEI DSSMSRKCKPLQCLSDIBWRIGHLAFFTLLVSTESDEL DSYMSRKCKPLQLSDIBWRIGHLAFFTLLVSTENDLAFLVDMTM HLMALNISLQGHSQIVTOMYDLRAFLAKLCL WETTELTRNILAFFTLLLVSTENDLAFLVDMTM HLMALNISLQGHSQIVTOMYDLRAFLAKLCL WETTELTRNILAFFTLLLVSTENDLAFLVDMTM HLMALNISLQGHSQIVTOMYDLRAFLAKLCL WETTELTRNILAFFTTLLVSTENDLAFLTVBNCTTLAVTSWT ASAPVQCLLLMAATFSPQGLAKPHSGTTPTC CFANNTKIPTGRLSTTATNICQFKEAVM  848 2198 A 6985 3 289 SVQYLFGRFTTHASTDGALAMKFTPLLST  SAPPOCLLLMAATFSPQGLAKPHSGTTPTC CFANNTKIPTGRLSTTATNICCHLLVVMS YDLVVAICHPTTLLSTOTTLGISLDSRLHAP MYFELSHLAVVDLAVACNTVPRMLVNLLIHP AKPISAGRAMQTTFSFTAVTECLLLVVMS YDLVVAICHPTFVLLIMMTTWCCTTLAVTSWT TGVLLSHHUNLLLPLPCPCPKYMHTPCELLA VLKLACADTHENEMVLAGAISGLVGPLSTIV	ł	1	1	ì	ł		
POLLNFGEFKCSDSDKKCHLWTAMDLGFFV VMGVIGGLIGATFNCLNKILAFYRARNVEN VMGVIGGLIGATFNCLNKILAFYRARNVEN KPKLVRVLESLLVSI,VTTVVVFVASMVLGEC RQMSSSQIGNDSPQLQVTENDSSIKTFEPD NDTYNDMATLFFNPQESALLQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSQLFVPSLLC GAAFGRLVANVLKSYGLGHPVSLLC GAAFGRLVANVLKSYGLGHPVGLRGVPLLEW LAWGKWTGDFFNKGNDHIVGLRGVPLLEW ETEVEMDKLRASDMEPNLTVVYPHTRIGSLV SILKTTVHHAFPVTENRGRESFMKROQLIS NNIKFKKSSLTRAGEGRKRSQSMKSYPSSEL RNMCDEHASEPAREDLLQMERRYTPY PRLYPPQSPEEDWTMEERFRLTFHGLILRSQ LVTILVRGVCYSESQSSAGPRLSVAEMAED YPRYPDHDLDLTLNPFMUVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGITTRHLTYPELQARLRGHYQTI  847 2197 A 6951 3 1994 NTMSSSVTNSAAGVEBLNIVQVTVPDNEERR LSSIEKIKQLREQVNDLFSRKFGBAIGVDFPVK VPYRKITENGCVVIDGMPGVVFKAPGYLEI SSMRRILRAAERIKTVIRPLPELSNGEYST VGRKKDDGGRVPGEKWERAYPFVEVONIST CLICKRSMSVSKEYNLRRHYOTNINSKHYDQV MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQVEDLAGNLWEKLR EKIRSFVAYSJAIDETIDNNT1QLAFFRGVDE NFDVSEBLLDTVPMTGTKSGNEIFSRVEKSIK NFCINWSKLVSVASTGTPPMVDANNGLVTKL SKRVATFCKGAELKSICCHIPESLCAQKKLMM DHVMDVVKSVSTGTPMVDANNGLVTKL KSRVATFCKGAELKSICCHIPESLCAQKKLMM DHVMDVVVKSVSTGTPMVDANNGLVTKL NSCHWSLVSVSTGTTPMVDANNGLVTKL KSRVATFCKGAELKSICCHIPESLCAQKKLMM DHVMDVVVKSVSTGTTPMVTLALFUTRUTH LINALNISLQGHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHATM LINALHSLAGCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHATLAVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHATLAVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHALVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHALVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHALVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHALVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHALVDMTM HLNALNISLQCHSQIVTOMYDLRAELAKLCL WETFILTENLAHPTLALVSHALVDMTM HLNALNISLQCHSQUTSTTLTCC CPANATNEPIPGLAEVSTRITOCPCKLAVM DVWGCRCHANTKFTPGLAXPHTVCC CPANATNEPIPGLASDMANTGLISTGTLUCPCREAVM HLNALNISLQCHSLSTVITCDCFKALVM HLNALNISLQCHSLSTVITCDCFKALVM HLNALNISLQCHSLSTVITCDCFKALVM HLNALNISLQCHSLSTVITCDCFKALVM HLNALNISLQCHSLSTVITCDCFKALVM HLNALNISLQCHSQUTSTTLATTTAVTECLLLVVMS YULVALCHEPTLYVTLLANGTLUCTLAT	1	1			,		1
KPKL VRVLESILIVSL VITVVVPVASMVLGEC RQMSSSQIGNDSFQLOVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESALQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFYPSLLC GAAFGRLVANVLKSYIGLGHYSGTFALIGAA AFLGGVRMTISLTVILIESTNETTYCLFINNTT LMYGKWTGDFFNGKIDVGLGRUPLLEW ETEVEMDKLRASDIMEPNLTYVVPHTRIQSLV SILRTTVHARPPVYTENKEREFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHLASEEPAKEDLLQPMLERRYTFY PNLYPPQSSEDWIMEERFPLTFHGLILLSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIBLDLDLTLLINPRIMYDVTPYMNPSFF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGITERHNLTYFEI, OALBOHYOTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEALGVDFPVK VPYRKITENFOCVVIDGMPGVVFRAPGYLEI SSMRRILBEAAFFIKFTVIRFLGLELSKGEYST VGRKKIDGGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTHISKHYDGY MERMEDELHELKIGLREVLGLISDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKRSFVAYSIADEITDINNTTQLAFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCNWSKLVSVASTGTPFNVDANNGLVTKL KSRVATFCRGAELKSICLHIPESLCAQKKLM DHVMDVVKSVNWICSRGNHSEFTLLYFEL DSQYGSLLYTEIK WLSRGLVLKRFFESLEIE DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPPTLKIVSRNESDGLNYP KIAELKTEFQKRLSDFKLYESSLTLTSSPFSTKI DSVHEELQMBVDLQCDTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDOWDSVLHAT  848 2198 A 6985 3 289 SVQYLFGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATTSPQCLAKPHSGTTITC CFNAINTKIPQGLESSTNICQPKEAVM MYFELSHLAVDIAYACNTVPRMLVNLLIP AKPISFRAGMMQTFLFSTFAVTECLLLVVMS YDLYVALCHELRYLAMFVRVCITLAVTSWT TGVLLSLHHLVLLIP,PFCRPQKIYSHFCELLL VMLS YDLYVALCHELRYLAMFTRLUSLUSLSSLHAP MYFELSHLAVOLIAYACNTVPRMLVNLLIP AKPISFRAGMMQTFLFSTFLAVTECLLLVVMS YDLYVALCHELRYLAMFTVRCCTLAVTSWT TGVLLSLHHLVLLIP,PFCRPQKIYSHFCELLL VMLS YDLYVALCHELRYLAMFTRCHLLIVVNSU TGVLLSLHHLVLLIP,PFCRPQKIYSHFCELLL VMLS YDLYVALCHELRYLAMFTRCHLLIVVNSU TGVLLSLHHLVLLIP,PFCRPQKIYSHFCELLL VLAKA	1		1				1
ROMSSSQIGNDSFQLQVTEDVNSSIKTFFCP MDTYNDMATLFPNQPGSLQLPHQDGTSF9V TLALFFVLYFLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHTSGTFLAIGAA AFLGGVVRMTSLIVILLESTNEITYGLPINAT LLMVGKWTGDFFNKGIVDHVGLRGVPLLEW ETEVEMDKLRASDIMEPNLTVVYPHTRIQSLV SILRITVHHAFPVVTENRGNEEPKKGNQLIS NNIKFKKSSLITAGEQRKRASOSMKSYPSSEL RNMCDEHLASEPPAEKEDLLQOMLERRYTPY PNLYPDQSPSEDWTMEEFRPLTFHGLILRSQ LVTLLVRQVCYSESQSSASQFRLSYAEMAED YPR YPDHDLDLTLLNFRMIDVDTYPMPSFF TVSPNTHVSQVFNLFRIMGLRHLPVVNAVGE IVGIITHNLTVEFLQARLRQHYQTI SSFKIKQLREQVNDLFSRKFGBAIGVDFPVX VPRKITTIPNGCVVIDGMPVVFKAPGVLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGRRKIDQEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHYQTHVHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSTTQSSVGKEYNLRHYQTHVHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSTTQSSVGKEYNLRHYGTHVHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSTTQSSVGKEYNLRHYGTHSFRYDEKSLK NFCINNSKLVSVASTGTPFMVDANNGLVTKL KSRYATFCKAELKSICGHPESLCAGNLKLM CHYMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFSLEEI DSFMSSRGKFPQLSISDWIRDLAFLAKLCL WETHLTRINLAHFPTLKIVKSNNESDGLNYP KIAELKTEFQKRLSDFKLYESSELTLSSPFSTKK DSVHEELQMBVDLQCNYLKTKSYDKVTGEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVHEELQMBVDLQCNYLKTKYDKVGTED FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWSVLHIAT ASAPVQCLLLMAATTSPQGLAKPHSGTIPTC CFNAINTKIPQCRESTWINGCPKEAVM MYFELSHLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP AKPISSELLAVVDIAYACNTVPRMLVNILLIP	1.						VMGVIGGLLGATFNCLNKRLAKYRMRNVHP
NDTYNDMATLFINGESALQLIFHQDGTFSPV TLALFFV/TYLLACWYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA AFLGGVVRMTISLTVILESTNEITYGLPINVT LMVGKWTGDFFNKGIYGHVGLGGVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTIVHHAFPVVTENGKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSNKSYPSSEL RNMCDEHJASEPARENQLQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDHDLDLTLLNPRMIVDVTPYNMPSFF TVSSNTHVSQVFNLFRTMGLRHLPVVNAVGE VGHTRINLTYFELQARIQHYQTI RTNSSSVTNSAAGVEDLNIVQVTYPDNEER LSSIEKIKQLREQVADLESKYGEAGVDFPVK VPYNKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRILLEAAEFIKFTVIRPLFGELSINGEVST VGKRKIDGCGRVPGEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTHSKHYDQY MERMREKLHELKKGLRYYLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINTTQLAFFIRGVDE NFDVSEBLLDTVPMTGTKSGNEIFSRVEKSLK NFCNWSKLVSVASTGTPFMVDANNGLVTKL KSRVATFCKGAELKSICCHFPSSLCAQKILKM DHVMDVVVKSVNWICSGILNHSEFTTLLYEL DSGYGSLLYTFIKKVSRGLVLKRFFFSLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTOMYDLIRAFLAKLCL WETHLTRNNLAHFFITLLYSFFSTLEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTOMYDLIRAFLAKLCL WETHLTRNNLAHFFITLLYSFFSTLEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTOMYDLIRAFLAKLCL WETHLTRNNLAHFFITLKVSRNESDGLNYPP KIABLKTEFQKRLSDFKLYESELILESSFFSTKL DSVHEELDMEVDLQGLBSQWDSVLHAT SVHEELGMEVDLQGLBSQWDSVLHAT SVHEELGMEVDLQGLBSQWDSVLHAT SVHEELGMEVDLQGLBSQWDSVLHAT SVHEELGMEVDLQGLBSQWDSVLHAT SVHEELGMEVDLQGLBSQWDSVLHAT ASAPVQCLLLMAATFSPQGLAKRHSGTTTICG CFNANTKPIPQRLESYTRITNIQCYEEAVM MYFFLSHLAVVDLAVACNTVPRMLVNLLHP AKPISFAGRAMQFFLFSTTRAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHLVLLLLPFCRPQKFRLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHLVLLLPFCRPQKFRLLVVMS	[		1	1		ì	KPKLVRVLESLLVSLVTTVVVFVASMVLGEC
TLALFFVLYFILACWTYGISVPSGLFVPSLLG GAFGRUAVNLKSSIGLGHIYSGIFALIGA AFLGGVVRMTISLTVILIESTNEITYGLPIMVT LMVGKWTGDFFNKGRYDHVGLGAVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIGSLV SILRTTVHHAPPVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEGRKRSGSMKSYPSSEL RNMCDEHIASEEPAEKEDILQOMLERRYTFY PNLYPDQSPSEDWTMEERFFLTHGLILRSQ LVTLLVRGVYSSGSSAGRLSYAEMAED YPRYPDIHDLDLTLNPRMIVDVTYPMNPSFF TVSPNTHYSQVFLKETRIGGLRHLPVVNAVGE IVGITTRINLTYEFLQARLRQHYOTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPMNPSFF SMRRILEAAFEIKTVHJQVTLAVENDEV SMRRILEAAFEIKTVHJGLELSNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGLESNGEVST VGKRKIDOEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNHFJGGTESTVLAGNLWELLR EKRSFVAYSIADIETIDNTTQLAFIRGGVE NFDVSEELLDTVPMTGTKSGNESFRVEKSLK NFCNWSKLVSVASTGTPMVDANNGLVTKL KSRVATFCKGAELKSICCHPESLCAQKLKM DHVMDVVVKSVNWICSRGPNDANALKLCL WETHLTRNNLAHFPILKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSFFSTKI DSVHEELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLSMFGSTTICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHAT SVEYELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLSMFGSTTICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHAT ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVQCLLLMAATSPQLAKPTSPJSKTK ASAPVGLLLMAATLGLISLDSSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLHIP AKPISFAGRAMOTTI-ERVATCHILGLISLDSSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLHIP AKPISFAGRAMOTTI-ERVATCHILLTUTSWT TGVLLSLHLVLLLLPFCRPQCKITV					l		
GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA AFLGGVVRMTISLTVILETNEITYGLPIMVT LMVGKWTGDFFNKGIYDIHVGLRGVPYLLEW ETEVEMDKLRASDIMEPNLTYVPHTRIQSLU SILRTIVHHAPPVVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEGRKRSGSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQOMLERRYTPY PNLYPDQSPSEDWTMEERRPLTHGLLRSQ LVTLLVRGVCYSESQSSASQRLSYREMAED YPRYPDIHDLDILLINPRMIVDVTPYMPSFFF TVSSPITHVSQVFNLFRIMGLRHLPVVNAVGE IVGIITRHNLTYEFLQARLRGHYQTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNVQVTVEPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYKRITHPIGCVULDGPFOVFKAFGVLEI SSMRILEAAEFIKFTVIRPLGELESNGEYST VGKRKIDQEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSEYSHLRHYQTNHSKHYDQY MERMRDEKLHELKGGLRYYLGLSDTECPE QKQVFANPSPTOKSPVQVVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPFMVDANNGLVTKL KSRVATFCKGAELKSICCHHPESSLCAQKKLKM DHYMDVVKSVNVINCSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYPP KIABLKTEPQKRLSDFKLYSSELTLFSSPFSTKI DSVHEELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLLSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHAT SOVHEELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLLSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHAT SAPPYQCLLLMAATFSPQGLAKPHSGTPTTC CFNANTKIPIQRLESYTRITNICQFKEAVM LFGLFSLFYVFTLLGRIFLGISLDSRLHAP MYFFLSHLAVVDIAYACNTVFRMLVNLLHIP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPFCAPQKIMFFCEILAV VKLACADTHNIEMWALAGAISGLVGPLSTIV	1				1		
AFLGGVVRMTISLTVILIESTINEITYGLPIMYT LMVGKWTGDFFNKGDIVIVGLGWGPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVYTENRGNEKERMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHLASBEPAEKEDLLQOMLERRYTPY PNLYPDGSPSEDWTMERPRIPTHFGLILLSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLEFREPLTHFGLILLSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLEFREPLTHFGLILLSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLEFREPLTHFGLILLSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLEFREPLTHGLERLLSQDTFVK VPYNRITTPRGCVVIDGMPGVVFKAPGYLEI SSMRRILEAAEFIKFTVHPLPGLELSNGEVST VGKRKIDQEGRVPQEKWERAYFFVEVQNIST CLICKRSMSVSKEVNLAFRYOTHNSKHYDQY MERMRDEKLHELKKGLRYYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKRSSVAYSIADIETIDINNTTQLAFFRGVDE NFDVSESLLDTVPMTGTKSGNEBISRVEKSLK NFCNWSKLVSVASTGTPMVDANNGLVTKL KSRVATFCKGAELKSICCHPESLCAQKKLKM DHVMDVVVSSVMVUSQLNISSGETTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISQGHSQIVTMYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLKLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLLVSRNESDGLNYPP KLAFLKTEPGKRLSDFWYDLRAFLAKLCL WETHLTRNNLAHFFTLLYST LYGERSTON HERENTYTH VALHER BAS 2198 A 6985 3 289 SVQYLPGRFTTTHASTODAPLMLKFTFLPSKTK ASAPVQCLLLMAATTSPQGLAKPHSGTPTTC CFNANTKIPJQCKEAW MYFLSHLAVVDIAYACCTVPRMLVNLLH ARPESTA	1	1	1	İ			1
ILMVGKWTGDFFNKGNYDIHVGLRGVPILLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRITVHHAFPVVTPNTGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKKSQSMKSYPSSEL RMCDEHIASEEPAKEDILQOMLERRYTPY PRILYPDQSPSEDWTMEERFRPLTHGLILRSQ LVTLLVRGVCYSESQSASQPRLSYAEMAED YPRYPDHIDLDLTLINPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRINGLGRILPVVNAVGE IVGITRHNLTVEFLQARLRQHYOTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIRQLREQVNDLFSRKTGEAIGVDFVK VPYRKITENPGCVVDDMPPGVVFKAPGYLEI SSMRRILEAAFFIKFTVIRPLPGLEISNGEYST VGKRKDOEGRVPGVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NDVSEELLDTVPMTGTKSGNEBESVEVSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIHFESLCAQKLKM DEHWIDVVVKSVMICSRGLNSESFTTLLYEL DSGYGSLLYYTEIK WLSRGLVLKRFFESLEBI DSFMSSRGKPLPGLSBWMRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPTLKLVSRNESDGLNYP KIAELKTEFOKRLSBUWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPTLKLVSRNESDGLNYP KIAELKTEFOKRLSBUWRDLAFLYDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPTLKLVSRNESDGLNYP KIAELKTEFOKRLSBUWRDLAFLYDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPTLKLVSRNESDGLNYP KIAELKTEFOKRLSBUWRDLAFLYDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPTLKLVSRNESDGLNYP KIAELKTEFOKRLSBUWRDLAFLYDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPTLKLVSRNESSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRFTRTHASTDAPLMKFTPLFPSKTK ASAPVCILLIMAATFSPQGLAKPHSGTTPITCC CFNAINTKIPIQRLSSYTRITNIQCPKEAVM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLSTFTAVTECLLLVMS YDLYVALCHPLRYLAIMTWRVCITLAVTSWT TGWLLSLIHLVILLPLPFCRPKIYHFFCEILA VKLALACADTHINENNVLAGAISGLVOFLSTIV VKLALACADTHINENNVLAGAISGLVOFLSTIV VKLALACADTHINENNVLAGAISGLVOFLSTIV VKLALACADTHINENNVLAGAISGLVOFLSTIV		1	1			1	
ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVYTENGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQQMLERRYTPY PNLYPOGSPESDWIMEERFRPLTHFIGLIRSQ LVTLLVRQVCYSESQSSASQPRLSYAEMAED YPRYPDHDLDLTLLNPRMIVDVTPYMPSFF TVSPRTHYSQVPRLPTRINGLRHLEPVNAVGE IVGIITRINLTYEFLQARLRQHYQTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLRQVNDLFSRKFGEAIGVDFPVK VPYRKITFNFGCVVIDGMPFGVVFKAPGYLEI SSMRRILEAAEFIKTVPLPGELISNGEYST VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHYQTNEKHYDQY MERMFDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSLADBETDNNTTQLAFIRGYDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPFMDANNGLVTKL KSRVATFCKGAELKSICCIHPESLCAQKLKM DHYMDVVVKSVNWICSRGLNISEETTILLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHPTLKLVSRNESDGLNYPP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTK YDKVGIFE FYKYLWGSYPKYHCAKILSMRGSTYICEQ LFSIMKLSKTK YCSQLKDSQWDSVLHIAT 848 2198 A 6985 3 289 SVQYLPGRTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNANTKIPIQRLESYTRITNIQCFKEAVM LDFLCHRDMGDNTTSTEFLLLGFFVQPRTQM 849 2199 A 6999 963 5 LDFLCHRDMGDNTTSTEFLLLGFVQPVQPRTQM WYFFLSHLAVVDIAYACNTYPRMLVNLHP AKPISFAGRMMQTFLFSTHAVTECLLLVVMS YDLYVALCHPLRYLAIMTWRVCTILAVTSWT TGVLLSLIHLVVLLLPLPCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVOPLSTIV						1.	
SILRITYHHAPPVYTENRGKEKEPMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQQMLERRYTPY PNLYPDQSPSEDWIMEERFRFLTHGLLIRSQ LYTLLVRGVCYSESSASQPRLSYAEMAED YPRYPDHDLDLTLINPRMIVDVTPYMNPSFF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGITTRIHLTYEFLQARLRQHYQTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSEKIKQLREQVNJESKYGEAGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRILEAAEFIKFTVIRPLPGLELSNGEFYST VGKRKIDOGGRVFGKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDDY MERMRDEKLHELKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQVELKM DHYMDVVVKSVNWICSRGLNHSEFTTLLYEL DSGYGSLLYYTEIKWLSRGLVLKRFFESLEBI DSFMSSRGKPLPGLSDWIPDLAFLAKLCL WETHLTRINLAHFPTLKLVSRNESDGLNYPP KIAELKTEFQKRLSDFWLTPSSTYTICED LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRFTRTHASTDAPLMKFTPLPSKTK ASAPVQCLLLMAATFSPQLAKPHSGTTPITCC CFNAINTKIPIQRLSSYTRITNIQCFKEAVM 1LFGLFSLFVYFTLLGMGTTLGLISLDSRLHAP MYFFLSHLAVVDIAYAGATSPQLAKFGTPFIPSTK CRNAINTKIPIQRLSSYTRITNIQCFKEAVM 1LFGLFSLFYVFTLLGMGTTLGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTTAVTECLLLVVMS YDLYVALCHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHLVLLLPLPFCRPKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	1	l	Ī	1		İ	1
RNMCDEHLASEPPAKEDLLQQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ LVTILLVRGVCYSEQSSASQPRLSYAEMAED YPRYPDIDLDLTLLNPRMIVDVTPYNNPSFF TVSPNTHVSQVFLKETMGLRHLPVVNAVGE IVGIITRHNLTYEFLQARLRQHYQTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTYPDNEKER LSSIEKIKQLREQNDLFSRKFEALGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILRAAEFIKFTVIRFLPGLELSNGEYST VGKRKIDQEGREVPGEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERNRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSTORSVQVPVEDLAGRILWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIHPESLCAQKILKM DHVMDVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYVTEIKWLSRGLVLKRFFSLEBI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HINALNISLQGHOVITQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLVSELTLFSSPFSTKI DSVHEELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHAT SVQYLFGRPTITHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPTTC CFNAINTKIPJQRLESYTRITNIQCPKEAVM LIFGLFREMGDNITSITEFLLLGPPYGPRIQM LLFGLFREDMGDNITSITEFLLLGPPYGPRIQM LLFGLFREDMGDNITSITEFLLLGPPYGPRIQM MYFFLSHLAVVDLAYACNTVPRIKLVNLHIP AKFISFAGRMMQTFLFSTFAVTECLLLVVMS YDLVVALCHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINEMMVLAGAISGLVGPLSTIV	}				Ì	1	
PNLYPDQSPSEDWIMEERRPLTHGLILRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDTLLNPRMVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVSIDITARINLTPEQARLRQHYQTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITHNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQGGRVPGEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRFYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKRSFVAYSIAIDEITDINNTTQLAFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCNWSKLVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHPPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPPSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLPGERTHASTDAVLMKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIQRLSYTRITNIQCPKEAVM MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLISTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSIHLLVLLLPLPFCRPKIYHFFCEILA VUKLACADTHINENMVLAGAISGLVGPLSTIV	1		İ	1		1	NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL
LVTLLVROVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGIITRINLTVEFLQARLRQHYQDT  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFFVK VPYRKITFNBGCVVIDGMPPGVVFKAPGYLBI SSMRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRYPQEKWERAYTFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRRYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFFDVSEBLLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIHFESLCAQIKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLRAFLAKLCL WETHLTRINLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSVPKYKHHCAKLLSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSVPKYKHHCAKLLSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT ASAPVQCLLLMAATFSPQGLAKPHSGTTPITC CFNAINTKIPIQRESYTRITNIQCPKEAVM MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLESTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHILVLLLPLPFCRPKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	1.				İ		
YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSFF   TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE   IVGIITARINLTYEFLQARLRQHYQTI     847   2197   A   6951   3   1994   NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER   LSSIEKIKQLREQVNDLFSRKFGEAIGVDFFVK   VPYRKITFINFGCVVIDGMPPGVVFKAPGYLEI   SSMRILEAAEFIKFTVIPLPGLELSNGEYST   VGKRKIDQEGRVFQEKWERAYFFVEVQNIST   CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY   MERMRDEKLHELKKGLRKYLLGLSDTECPE   QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR   EKKRSFVAYSIADETDINNTTQLAFIFRGVDE   NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK   NFCNWSKLVSVASTGTPPMVDANNGLVTKL   KSRVATFCKAEKSICCIHPESLCAQKLKM   DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL   DSQYGSLLYYTEIKWLSRGLVKKFFESLEEI   DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM   HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL   WETHLTRINLAHFPTLKLVSRNESDGLNYIP   KIAELKTEFGLRSDFKLYESELTLFSSPPSTKI   DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE   FYKYLWGSYPKYKHCAKILSMFGSTYICEQ   LFSIMKLSKTKYCSQLKDSQMDSVLHAT   SAPVQCLLLMAATFSPQGLAKPHSGTTPITC   CFNAINTKIPJCRLESYTRITNIQCPKEAVM   SASPVQCLLLMAATFSPQGLAKPHSGTTPITC   CFNAINTKIPJCRLESYTRITNIQCPKEAVM   MYFFLSILAVVDIAYACNTVPRMLVNLLHP   AKPISFAGRMQTFLFSTFAVTECLLLVVMS   YDLYVAICHFLRYLAIMTWRVCITLAVTSWT   TGVLLSIHLVLLLPLPFCRPQKIYHFFCEILA   VLKLACADHINENMVLAGAISGLVGPLSTTV			1	1		i	
TVSPITHVSQVFNLFRTMGLRHLPVVNAVGE IVGITTRINLTYEFLQARLRQHYQTI  847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILBAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVPGEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDERLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSTVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQKLKM DHYMDVVVKSVNWICSRGLNHSEFTTILLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRINLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELILFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLDGRIFTRIHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIGRLESYTRITNIQCPKEAVM LLFGLFSLFYVFTLLGNGTLGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLHP AKPISFAGRMMQITLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGL VGPLSTIV	1.			1	1		1 1
IVGITRHNLTYEFLQARLRQHYQTI	1	1	ì	1		ĺ	l
847 2197 A 6951 3 1994 NTNSSSVTNSAAGVEDLNIVQVIVPDNEKER LSSIEKIKQIREQVNDLFSRKFGEAIGVDFPVK VPYRKITENPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCHIPESLCAQKLKM DHYMDVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLRAFLAKLCL WETHLTRNNLAHPFTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPPSTKI DSVHEELQMEVDLQCNTVLKTK YDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLFGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPTC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LLFGLFSLFYVFTLLGNGTLGLISLDSRLHAP MYFILSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHLVLLPLPFCRPQKIYHFFCEILA VLKLACATHINENMVLAGAISGLVGPLSTIV	1					1	
LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITRNPGCVVICKMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRYYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLRAFLAKLCL WETHLTRILAHPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CPNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFVYFTLLGNGTILGLISLDSRLHAP MYFTLSSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	2/7	2197	Α	6951	3	1994	
VPYRKITENPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLEISNGEYST VGKRKIDOGRVPFOEKWERA YFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLISMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT 848 2198 A 6985 3 289 SVQYLPGRFTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNILHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKYTHFFCEILA	047	2191	1	0,51		1 1331	
VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSFTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCNWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNILAHPFTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTILFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAFVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LLFGLFSLFVVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	1		Ì				
CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEBLLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHILTRINLAHFPTLKLVSRNESDGILNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKLISMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMIKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLVAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWNCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	1.	1			1		SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST
MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEBLLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTILFSSPFSTKI DSVHEELQMEVIDLQCNTVLKKTYDKVGIPE FYKYLWGSVYKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNILHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWXCITLAVTSWT TGVILSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV		1	1	1	1		
QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM ASAPVQCLLLMAATFSPQGLAKPHSGTIPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM MYFFLSHLVAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGYLLSLHILVLLLPPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV		ĺ		ļ			· · · · · · · · · · · · · · · · · · ·
EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCNWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQVKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT S48 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQOLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLIPLFFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV							1
NFDVSEELLDTVPMTGTKSGNÈIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATTCKGAELKSICCIIHPESLCAQWLKM DHVMDVVVKSVNWICSRGLNHSEPTTILYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATTSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLVAVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV				1	i		
NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATTSPQGLAKPHSGTIPITC CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLVAVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLPLPFCRPQKIYHFFCELLA VLKLACADTHINENMVLAGAISGLVGPLSTIV		1	1		1	1	
KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLGFPVGPRIQM LLFGLFSLFLVFVFTLLGNGTLGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV		1	1		<b>\</b> .		
B48 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFLAVVLLHP AKPISFAGRMAQTFLFSTAVTECLLVVMS YDLYVALCHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV		1	ł				
B48 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLHLAVVDLLHP AKPISFAGRMMQTFLFSTAVTECLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV			1		1		
HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM LFGLFSLFVFTLLGFPVGPRIQM LLFGLFSLFVVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV		1	{		1		
WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVILSLIHLVLLIPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV			1				
KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV						1	
BASE 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFVFTLLGNGTLGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	Ì		1				
FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	1		1	1		}	
LFSIMKLSKTKYCSQLKDSQWDSVLHIAT  848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV					1	[	
848 2198 A 6985 3 289 SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFVVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	1			1	1	}	
ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM  849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFVVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	848	2198	A	6985	3	289	
849 2199 A 6999 963 5 LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV					1		
LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV					<u> </u>		
MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV	849	2199	A	6999	963	3	
AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV			ļ	1	1		
YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV			1		1		
TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV		1	1		[	1	· ·
VLKLACADTHINENMVLAGAISGLVGPLSTIV			ļ				- 7
			İ				
			1		1	1	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG KVARRRVGATWLLHLAVADLLCCLSLPILAV PIARGGHWPYGAVGCRALPSIILLTMYASVLL LAALSADLCFLALGPA WCLRFS/GACGVQVA CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA SCHSALLCWAARRCRPLGTAIVVGFFVCWAP
				<u>.</u>		YHLLGLVLTVAAPNSALLARALRAEPLIVGL ALAHSCLNPMLFLYFGRAQLRRSLPAACHW ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	A	7011	1	2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI SRGVLVCDECCSVHRSLGRHISIVKHLRHSA WPPTLLQMVHTLASNGANSIWEHSLLDPAQV QSGPALKQTPKDKVVHPIKSEFIRAKYQMLAF VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE TCLRLLSLGAQANFHPEKGTTPLHVAAKAG QTLQAELLVVYGADPGSPDVNGRTPIDYARQ AGHHELAERLVECQYELTDRLAFYLCGRKPD HKNGHYIIPQMADSLDLSELAKAAKKKLQAL SNRLFEELAMDVYDEVDRRENDAVWLATQN HSTLVTERSAVPFLPVNPEYSATRNQGRQKL ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD NLELSLRSQSDLDDQHDYDSVASDEDTDQEP LRSTGATRSNRARSMDSSDLSDGAVTLQEYL ELKKALATSEAKVQQLMKVNSSLSDELRRLQ REIHKLQAENLQLRQPPGPVPTPPLPSERAEH TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL YRIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK LSRHGSGADSDYENTQSGDPLLGLEGKRFLE LGKEEDFHPELESLDGDLDPGLPSTEDVILKT EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA VTEMASLFPKRPALEFVRSSLRLLNASAYRLQ SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA KAAKQLVTITTREKKQ
852	2202	Ā	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL TVKGLLKPSFSPRNYKALSEVQGWKQRMAA KELARQNMDLGFKLLKKLAFYNPGRNIFLSP LSISTAFSMLCLGAQDSTLDEIKQGFNFKMP EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID QRLQPQRKFLEDAKNFYSAETILTNFQNLEM AQKQINDFI/ESKTHGKINNLIENIDPGTVMLL ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL LSRRVVDSVPRLHMTGTFDLKKTLSYIGVS KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM DERGTEGAAGTGAQTLPMETPLVVKIDKPYL LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1	3293	MTHACNPSTLGGGGRRITRSHGRRRSSRGPV ARHVAAGAGHENKHGGSRRFPAGVAPRRAM ANVSKKVSWSGRDRDDEEAAPLLRRTARPG GGTPLLNGAGPGAARQSPRSALFRVGHMSSV ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY ESLDYDNSENQLFLEEERRINHTAFRTVEIKR WVICALIGILTGLVACFIDIVVENLAGLKYRVI KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

CECTO	SEC ID	Me+	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	Seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Laucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i	]		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ì		peptide	-	/=possible nucleotide deletion, \=possible
	1	j	İ	sequence	ļ <u>.</u>	nucleotide insertion
						LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH
		ì	į			VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI
	l	l	ł	ĺ	İ	HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE
		ł	l		į	KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG
			İ	ĺ	}	ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG
						NMWDLSSPGLINFGRFDSEKMAYTHEIPVFI
	ľ	1	İ	ĺ	1	AMGVVGGVLGAVFNALNYWLTMFRIRYIHR
	1		1	ļ		PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL
		ŀ		İ	1	QGGSMŞYPLQLFCADGEYNSMAAAFFNTPEK
		ļ	j	]	J	SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG
l	-					AAIWADPGKYALMGAAAOLGGIVRMTLSLT
		ł				VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE
1	1	1	1	1		GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV
			1			MSTPVTCLRRREKVGVIVDVLSDTASNHNGF
	İ	l				PVVEHADDTOPARLOGLILRSQLIVLLKHKVF
		1				VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH
	1	1				VSQDERECTMDLSEFMNPSPYTVPQEASLPR
]	1	l'		1	,	VFKLFRALGLRHLVVVDNRNQVVGLVTRKD
)		}	}	l		LARYRLGKRGLEELSLAQTGPKAQATAEGRV
·	1	ļ		1		AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP
	1	1	Ì			LSLEELSERYESSHPTSTASVPEQDTAKHWNQ
				·	į	LEQWVVELQAEVACLREHKQRCERATRSLL
ļ		1	}	1	ł	RELLQVRARVQLQGSELRQLQQEARPAAQAP
		İ	1			EKEAPEFSGLQNQMQALDKRLVEVREALTRL
	1	1	1	ļ		RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ
İ		[	[	1		QEEQGREVACGALQKNQEDSSRRVDLEVAR M
064	2204	<del>  </del>	7037	139	2604	AGTWEPRPYDOAKETGAPGSQPPVPPMELRP
854	2204	A	/03/	139	2004	WLLWVVAATGTLVLLAADAQGQKVFTNTW
	}			j	1	AVRIPGGPAVANSVARKHGFLNLGQIFGDYY
Ì	1	1		1	ļ	HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL
						EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG
	ĺ	1	1	1		VTO\RDLMVKAAWAQGYTGHGIVVSILDDGI
İ		ì		1		EKNHPDLAGNYDPGASFDVNDQDPDPQPRY
		1		1		TQMNDNRHGTRCAGEVAAVANNGVCGVGV
	1			i	J	AYNARIGGVRMLDGEVTDAVEARSLGLNPN
		1			ì ·	HIHIYSASWGPEDDGKTVDGPARLAEEAFFR
1				1		GVSQGRGGLGSIFVWASGNGGREHDSCNCD
ł	1	1		1	1	GYTNSIYTLSISSATQFGNVPWYSEACSSTLA
	i	1	1		1	TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS
1	Į	1	1			
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI
					1	APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS
						APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS
					,	APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD
Dec.	2005		7050	2		APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	1441	APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA
855	2205	A	7058	3	1441	APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGGGERTAFIKDQSAL  QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF
855	2205	A	7058	3	1441	APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL
855	2205	A	7058	3	1441	APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGGGERTAFIKDQSAL  QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS
				·		DYCRLLCAYILGNDFTDLFDIVITNALKPGFF SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFII DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206		7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHILLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLGLP WALIFFSFASGTFQLVVLYLFSITTSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	A	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	лисleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		J	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	! 	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1	1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	<del></del>	<del>                                     </del>	<del>                                     </del>	<del> </del>	<del> </del>	HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ
		1		1	Į	DHIYHPO*GSRGHHCPRPVPRPRLLGLGPSLP
	1	i	1	1	{	CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI
001		1		}		LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL
		1			'	KSAMILO .
862	2212	A	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF
002	22.12	^	/	1000	1	YKNIOKLSFSNYVYHQNYVFSSDWSYDF
863	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP
003	2213	^	1212	124	1275	LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA
	ì	1	1	1	1	RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ
		1	1		1	CIKPNYQSPPKECDYNILANSVA
000	2214	-	7214	845	1619	SDKGGKKADRKNHLRHAFPLLPHRVRERLH
864	2214	A	1214	073	1 10.5	DPKVPVDADHVQGQDPGRAAHDIHGEDVTE
	İ					KVSKDPLAPDEVGDTDEGHDRHGHREVGQR
	İ	1		1		HGHDOEEVAYEERACEGGKFATVEVTDKPV
		1	ł	1	ł	DEALREAMPKVAKYAGGTNDKGIGMGMTV
		1	ŀ	Į.	}	PISFAVFPNEDGSLOKKLKVWFRIPNQFQSDP
		1	1 :	1	ŀ	PAPSDKSVKIEEREGITVYSMQFGGYAKEAD
	1	1	ŀ	į		YVAQATRLRAALEGTATYRGDIYFCTGYDPP
			İ		]	MKPYGRRNEIWLLKT
	<u> </u>		<del></del>	7.50	(02	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS
865	2215	Α	7246	559	682	
					1210	LANMAKPRLY TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL
866	2216	A	7257	641	1310	DLKKSDFSTRWQKQRCPVVKSKCRENASPFF
	1	1	1		ł	FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV
			1		1	
ı		1		1	ł	QIPLTESYCGPCPKNWICYKNNCYQFFDESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV
i		1	1		1	KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT
ì		1	1	1	1	
i.		l l		}	Į.	IIEMQKGDCALYASSFKGYIENCSTPNTYICM
				<del></del>	1006	QRTV SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI
867	2217	Α	7288	151	396	MPFFQTLWLMNANRFCSIFTTTNVANNCWW
		1	<b>1</b>		]	
						TPYHCWLSVVVCRCESHGI
868	2218	A	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP
	1	1	i			KGSCPAGGSRMVSESD*EGRGC*ASYPCAC*
	_	1				AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	Α	7332	1223	332 .	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR
		1			1	FLTLCTWLLLLGPGLLATVRAECSQDCATCS
ļ	į.			ļ	İ	YRLVRPADINFLACVMECEGKLPSLKIWETC
	ł	1	1	j	}	KELLQLSKPELPQDGTSTLRENSKPEESHLLA
		1				KRYGGFMKRYGGFMKKMDELYPMEPEEEA
	1	1	İ		1	NGSEILAKRYGGFMKKDAEEDDSLANSSDLL
	1	i	-	ł	{	KELLETGDNRERSHHQDGSDNEEEVSKRYGG
j	1	1	-			FMRGLKRSPQLKEKAKELQKRYGGFMRRVG
l	1			-	[	PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE
		1				SYSKEVPEMEKRYGGFMRF
870	2220	A	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG
1			1	1		AGQGRGREGAESGGSRGEGPGSDGRLPATGD
1			ł	1	1	FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT
1		ł	i			CONTRACT OF CONTRACT A AMERICAN DELL'ANDIA
				l		TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL
						KGLOLLLSLLAFICEEVVSQCTLCGGLYFFEF
						KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD
						KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD
						KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF
						KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT
271	2221	A	7403	3	393	KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA
871	2221	A	7403	3	393	KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod .	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Mcthionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG PFIC  FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC PGGS*PQATLHLDRMRVSASPTKEIQVKKYK CGLIKPCPANYFAFKICSGAANVVGPTMCFED RMIMSPVKNNVGRGLNIALVNGTTGAVLGQ KAFDMYSGDVMHLVKFLKEIPGGALVLVAS YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
						SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE GWPELLEMEGCMPPKPF
873	2223	A	7429	2242	2394	ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVQ DHPGQHCETPSLLKIERKLF
874	2224	A	7468	146	894	PCTSCVLWATLHLPASTRKAPQAECGMISITE WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE MSSLNLDHWLKGAKREEWEPPPQSPALTHSP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE AEMQTPRRLGWGWYHTLTLYLWEEK
875	2225	A	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG SERHEP*HGGVLFRLGPSAPPGKL
876	2226	A	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
877	2227	A	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC TFKDKVLVAARRNASAVVLYNEERYGNITLP MSHAGTGNIVVIMISYPKGREILELVQKGIPV TMTIGVGTRHVQEFISGQSVVPVAIAFITMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS
878	2228	A	7586	315	1232	ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG
879	2229	A	7605	479	391	TESWKLKWWSPTCLDQLNGSAPGNVFIHG
880	2230	A	7612	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL
881	2231	A	7615	291	1452	SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN
						SGKYATTARNSFIVLIIFTICFVPYHAFRFIYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIQSSSKST
882	2232	A	7617	67	379	RQMALLKANKDLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	Α	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK AVATVGPISVAVGASHVFFQFYKKGKHLSS
884	2234	A	7638	2640	2861	APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP GLSCHTSHSG
885	2235	A	7642	201	455	PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRONRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A .	7702	242	1298	APSHRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707	185	2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKILINPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETTSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP DESTKTKDQILTSRINAVERDLLEPSPADQLG NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV TREPARRLFLFGEEPSKLDQDVLAALECADV DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
\$90   2240   A   7711   360   269   RHMPVIPALWEAEVGGLLEPRSSISAWATE	•	İ			1		
\$1   \$241   A   \$7721   \$61	800	2240	<u> </u>	7711	360	269	
VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL PAMPMQGGAQSPEELRAAVLQLRETVVQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPRDPGHVVEQLSRSLQTLK DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ LERTVVQQKETLASARAIRELTGKLARCEGL AGGKARGAGATGKDTMGDLPRDPGHVVEQ LSRSLQTLKDRLESLEHQLRANVSNAGLPGD FREVLQQRLGELRQLLRKGAELEDEKSLLH NETSAHRQKTESTLNALLQRVTELERGNSAF KSPNAFKVSLPLRTNYLLYGKIKKTLPELYAFT ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE WGNNPIELLINDKVAQLPLFVSDGGKWHHICV TWTTRDGMWEAPQDGKLGTGENLAPWEPI KPGGVLILGQEQDIVGGRFDATQAFVGELSQ FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD NNVDVFGGASKWPVETCEERLLDL  893 2243 A 7729 3554 2419 LTAGTAMNYPLITLEMDLEDLFWELDRL DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIOLVLVLIEHRFQTRSSTET FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLACIAVDRYLAIV HAVHAYRHRALSHITTGGTIWLVGFLLALPEI LFKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRPAQRGKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAFICCLNPMLYTFACVFRSDLSRLLTKLG CTGPASLCQLPFSWRRSSLESESENATSLTTE  894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGALIPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D  895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR					<u> </u>		KLPWEPSFLIKMQIIRHSEQTLKTALISKNPVL VSQYEKLDAGEQRLMNEAFQPASDLFGPITL HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN KRSIYIQSIGSLGNTRIISEEVIKWLTGYCKAYF YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS WNFVFGQASLTDGVGIFSFARYGSDFYSMHY KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL EEADRRPLNLCPICLHKLQCAVGFSIVERYKALVRWIDDESSDTPGATPEHSHEDNGNLPKPV
DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG CTGPASLCQLFPSWRRSSLSESENATSLTTF  894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D  895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR LWLSLFLHAGKEAPHCPRTRPL	892	2242	A	7723	2		VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL PAMPMQGGAQSPEEELRAAVLQLRETVVQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPRDPGHVVEQLSRSLQTLK DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ LRETVVQQKETLASARAIRELTGKLARCEGL AGGKARGAGATGKDTMGDLPRDPGHVVEQ LSRSLQTLKDRLESLEHQLRANVSNAGLPGD FREVLQQRLGELERQLLRKGAELEDEKSLLH NETSAHRQKTESTLNALLQRVTELERGNSAF KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE WGNNPIELLINDKVAQLPLFVSDGKWHHICV TWTTRDGMWEAFQDGKKLGTGENLAPWHPI KPGGVLILGQEQDTVGGRFDATQAFVGELSQ FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD NNVDVFGGASKWPVETCEERLLDL
VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D  895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR LWLSLFLHAGKEAPHCPRTRPL	893	2243	A	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLACIAVDRYLAIV HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
LWLSLFLHAGKEAPHCPRTRPL	894	2244	A	7738	670	287	VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
	895	2245	A	7753	119	278	,
	896	2246	A	7754	1	372	LWLSLFLHAGKEAPHCPRTRPL SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- eotide seq- uence	SEQID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QVSTAALAVLLCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775 •	85	496	SCQTTQPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	A	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796		807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRPQKGTAASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

	000 10	37-4	1 000	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of nucl-	peptide	1104	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	!	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	neuce	]	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	ualoc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciice	1	Ì	717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
•				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ļ	peptide	304	/=possible nucleotide deletion, \=possible
	ł		1	sequence		nucleotide insertion
903	2253	A	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
700			''	-	1	VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
		l	İ	Ì	]	RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
				ŀ		ARLLYESRKRGMLENCILLSLFAKEHLOHMT
			İ			EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
		į				EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
				Ì		LFEKPR
904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG
* * *				1	1	AGARLTGWTMNVFRILGDLSHLLAMILLLGK
ļ	ļ		1 .			IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
	Ì	l			1	ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF
						DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
ı i		l	1			WTFSIYLESVAILPQLFMISKTGEAETITTHYL
	1 .	i				FFLGLYRALYLANWIRRYQTENFYDQIAVVS
l .	ľ	l				GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL
						RSYSSI
905	2255	A	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA
		1	į			QTEMVRTLERKLEAKMIKEESDYHDLESVVQ
1		1	,			QVEQNLELMTKRAVKAENHVVKLKQEISLL
ľ			1 .	ĺ		QAQVSNFQRENEALRCGQGASLTVVKQNAD
			1			VALQNLRVVMNSAQASIEQLVSGAETLNLVA
					<u> </u>	EILKSIDRISEVKDEEEDS
906	2256	A	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
	i					LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
1				ŀ	ĺ	HQPQTGSGESSGASGDKDHLYSTVCKPRSPK
					į	PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA
			1	ļ.		TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
			1		ĺ	RPSLPSSPSPGLPKASATSATLELDRLMASLSD
1		Ì			Į.	FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
		ĺ				KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
1		1		1	1	NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
Ì		i	1	1		GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI
		ļ	1			RHKMVTALGTHWHPEHFCCVSCGEPFGDEG
			1	)	1	FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN
į		1				YISALSALWHPDCFVCRECFAPFSGGSFFEHE
		1				GRPLCENHFHARRGSLCATCGLPVTGRCVSA
	1	1			1	LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
	10055	1	7000	1702	1671	CQPCFLKLFG
907	2257	A	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
000	10050	<del> </del>	7040	110	1170	YCKSQAWG KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
908	2258	A	7842	110	1172	
1				1	1	SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET
		1		1		YTCPKMIEMEQAEAQLAELDLLASMFPGENE
	1		[		1	LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
			1		1 .	TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV
	1		1			CILNATEWVREHASGYVSRDTSSSPTTGSTVQ
			1		1	SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
1	1		1			SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
	[	1				SLSOFSNIPGRPGVVCVEGPQSACEEF WARLK KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
1		1				EEKVFSVNGARGNHMDFGQLYQFLNTKGCG
		1	1		1	DVFQMFLWV
000	2250	<del>- </del>	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
909	2259	A	1810	300/	2723	LISSEILLIPSKYLFESK
010	2260	<del>  </del>	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP
910	2260	A	/684	412	40/4	PSHRVNAEPGCVVTNACASGPCPPHANCRDL
1	1			1		WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
1	-	ļ	1	1		SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD
1		i		1		QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK
1.			_1			A SAN WAR I COL CHOD ANY OLD LUCHY

mucle- duide seq- uence    Peptide	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Sequence   Sequence			100				F=Phenylalanine, G=Glycine, H=Histidine,
uence    09/496   correspondi   pist amino acid residue of peptide sequence   peptide sequence   peptide sequence   peptide   sequence   peptide		F - F -				corresponding	I=Isoleucine, K=Lysine, L=Leucine,
mence and provided a mino said of residue of peptide residue of peptide sequence peptide sequence and peptide sequence peptide sequence se				09/496	correspondi	to last amino	
residue of peptide sequence   Y-Tyrosine, X-Unknown, "-stop codon,			l i	914	ng to first		
peptide sequence  /possible nucleotide insertion  TNGC/HCKEFHYRPRGDSCLP/CDCYPYGST SRSCAPHSGC/PCPCPGALGR/CNSCDSPFAEV TASGCRYLYDACPKSLRSGVWWPOTKEGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFREISLLLDCLELINKTALDTME AKKLAGRLEVTGHTDHYFSQD/WYTARIL AHLAFESHQQGFGLTATQDAHNEINLLWA GSALLAPETGDLWAALGQRAPGGSPGSGALV RHLEEVAATLARNMELTYLNPMGLVTPNML SIDRMEHPSSFRGARRYRRYHSNLFRGQDAW DPHITVLLPSGSPRSPSSEVLSSSISMSTTSS VYPPAPPFEPPGISIIILLVYRTLGGLLPAQFQ AERGARLPONYWMSPWYSWYHGRINIR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHOWWTARDCELI VHRNSKAICVSWTPGTHAR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHOWWTARDCELI VHRNSKAICVSWTPGTHAR GILESPISLEFRLLQTANRSKAICVQWDPFGLA EQHOWWTARDCELI VHRNSKAICCSRTGT PGVLMDASPRERLEGDLELLAVFTHVVAVS VAALVLTAAILLSLRSINNYRGHANVAAA LGVAELFILLGHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVERRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVLEPLIWSFAGEVVLVIVMNGTMFLIA ARTSCSTGQREAKKTSALTRSSFLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPGCRKAPEEAR PAPGLGFGAYNNTALFEESGLIRITIGASTVSS VSSARSGRTQDQDSGGRGSVLRDNVLVVHNGS AADHTDISLQAHAGFTDLDVAMFHDAGA DSDSDSLSLEERRSLSTSSEDGRTGRGF QRPLCRAAQSERLITHPKDVDGMDLTSKYDAN NNQPDPALTSGDETSLGRAQRGGRIKNRL QYPLVPQTRGAPELSWCRAAATLGFRAVPAAS YGRIYAGGGGSLSQPASRYSSREGLDLLERR QLSSERLEEAPAPVLREISRPGSGCCMDAPG RLEFKDRGSTLPRRQPFTDVTGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPGDFR DTQALLSATQMDLRRRDVTVQGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPGPFLP LSPQCLSEBDLISSSSNSSREGLDLLERR QLSSERLERAPAPVLRESRPGSGCCMDAPG RLEFKDRGSTLPFRRQPFTDVTGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPGSPER DTQALSATQMDLRRRDVTVQCSRARAYAL LLQHPVLVWPRYPVDVGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPGSFR DTQALSATQMDLRRRDVTVQCSRARAYAL LLQHPVLVWPRYPVDVGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPGSFR DTQALSATQMDLRRRDVTVQCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSLARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPRYPVPCUCCSRARAYAL LLQHPVLVWPTNYNCKELLERGTTYTL LVSQLRAQGEGCVQFQAUSAGGFFASTQCV VAPLRADGDKPRAHLTVVRQTFTQHERGPTTYNCHELDPSGD ALMWE							
Inucleotide insertion						sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
TNGCHKEFHYRPRGSDSCLPCDCYPGST SRSCAPHSGCPCPCPGALGRCONSCDSPAEV TASSCRVIJDACPKSLRSGVWWPQTKFGVL ATVPCPRGALGLRGAGAVRLCDEAGQWLE PDLFNCTSPAFREISLLLDGLEINKTALDTME AKKLAGRLREVTGHTDHYFSQDVRVTARIL AHLAFESHGQOFGLTATQDAHNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSGALV RHLEEVAATLARNMELTYLPMGLVTPNML SDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPPSPSEVLPTSSSEINSTTSS VVPPAPPEEPGISIIILLVYRTLGGLLPAGPQ AERGARLPQNPVMNSPVVSVAVHGRRFLR GILESPISLEFRLQTANRSKAUCVGWDPGLA EQHGVWTARDCELVHRNGSHARCRSTRGT FGVLMDASPRELEGDLELLAVFTHVVVAVS VAALVLTAAILLSIRSIKSNVRGHANNVAAA LGWAELHFLIGHRITHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR PYHALGWGVPAVLLGLAVGLDVGUVGHPGLA ARTSCSTGQREAKKTSALTIRSSFILLLUSA SWLFGLLAVNHSLAFHLYHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGFGG,YNNTIALFEESGLTGLGGCGLGAV SVSARSGRTQDQDSQRGRSYLRNDVLVHMOS AADHTDHSLQAHAGPTIDLVAMPHDAGA ADDSDSDLSLEERESLSPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGONLYWHGS AADHTDHSLQAHAGPTIDLVAMPHDAGA ADDFDALTSGDETSLGRAGRQRKGRILINGL QYPLVPQTRGAFELSWCRAATLGGRRAYPAAS VGRYAGGGTGSLSQPASRYSSREQLDLLRRQ QUSBCRLEERAPAVLRILSPACATLGGRRAYPAAS VGRYAGGGTGSLSQPASRYSSREQLDLLRRQ QUSBCRLEERAPAVLRILSPACATLGGRRAYPAAS VGRYAGGGTGSLSQPASRYSSREQLDLLRRQ QUSBCRLEERAPAVLRILSPACATLGGRRAYPAAS VGRYAGGGTGSLSQPASRYSSREQLDLLRRQ QUSBCRLEERAPAVLRILSPACDLPOPPLP LSPQQLSRDPLLPSRRDLSISRSSNSREQLDQ VPSRHFSREALGPLPQLRAREDSVSQPSHIGP STEQLDLSSLASSNSSRSJVGSSTFLGPHT TATPSATASVLGPSTPRSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLPSRRDLSTPSSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLSRRDLSTPSSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLSRRDLSTSPSSTEDDPSPR TATPSATASVLGPSTPRSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLSRRDLSTPSSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLSRRDLSTPSSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLSRRDLSTPSSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLPOPPLP LSPQRQLSRDPLLSRRDLSTPSSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLYVQLOTSTSTPVFIY TATPSATASVLGPSTPRSATSHSISSLSPSDEPS DTQALSATQAMDLRRCDLYVQLOTSTSTPVFIY TATPSATASVLGPSTPRSATSHSISSLSPSDEPS DTQALSATQAMDLRCDLYVQLOTSTSTPVFIY TATPSATASVLGGGFAPS		1			• •		
SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVI-TACAPCHSLRSGQWWPQTKTGVL ATVPCPRGALGLRGAGAVRLCDEAGOWLE PDLFNCTSPAPRELSLLLDGLEINKTALDTME AKKLAGRLREVTGHTDHYPSQDVRVTARLL AHLLAFESHQQGFGLTATQDAHNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSAGLV RHLEEVAATLARNMELTYLNPMGLVTPINML SIDRMEHPSSFGARRYPRYSINLFRGQDAW DPHTHVLLPSQSPRSPSEVLPTSSEISHTTSS VVPPPAPPEPEPGISIILLVYSTLTGGLLPAQFG AERGARLPQNPVMNSPVVSVAVPHGRNFLR GILESPISLERELLQTANRSKALVQWDPGLA EQHGWWTARDCELVHRNGSHARCRCSRTGT PGVLMAASFREALEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNYRGHANVAAA LGVAELLFLLGHRITHNQLVCTAVVLLHYFF LSTFAWLFVGGLHYRMQGVPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNDDF CWISVIEPLIWSRAGPVVLNVMGTMFLLA ARTSCSTGQREAKKTSALTLRSSFILLLUSA SWLFGLLAVNHSLAFFLYHAGLCGLOGLAV LLFCVLNADARAAWMPACLGRKAAPEEAR PAPDLGPGARVNTALFEESGLATTLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMPHADAGA DSDSDSISLEERSLSSPSSESEDNGRTGGRF QRFLCRAAQSERLLTHKDVDGNDLLSYWPA LGECAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGLINRL QYPLVPOTTRGDETSLGRAQRQRKGLINRL QYPLVPOTTRGDETSLGRAQRQRKGLINRL QYPLVPOTTRGDETSLGRAQRQRKGLINRL QYPLVPOTTRGDETSLGRAQRQRKGLINRL QYPLVPOTTRGDETSLGRAQRQRKGLINRL QYPLVPTTRGDETSLGRAQRQRKGLINRL QYPLVPTTRGDETSLGRAQRQRKGGSR DALDIGAPREWISTLPPPRRTRDLDTDPOPPLP LSPQGLSGDPLLFSRPLDSLSSRSNSREQLDD VPSRHPSREALGPLPQLLRAREDSVSGPSIGD STEDLDLSSLASSNSSASSSSTELGHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALSATQAMDLRRCPVULGFDLSGLSVA MQLPQGLAVALLAGLPVFGLYSSFYPFIY FLEFGTSRHSVESLCVPPGVVLGDLISGISVA MQLPQGLAVALLAGLPVFGLYSSFYPFIY FLEFGTSRHSVESLCVPPGVVLGDLISGISVA MQLPQGLAVALLAGLPFVGLYSSFYPFIY FLEFGTSRHSVESLCVPGVVVLQCSTARAYAL LLCHLPHVLWPRYPVRWULGDLISGISVA MQLPQGLAVALLAGLPFVGLYSSFYPFIY FLEFGTSRHSVESLCVPGVVVLQCSTARAYAL LLCHLPHVLWPRYPVRWULGDLISGISVA MQLPQGLAVALLAGLPFVGLYSSFYPFIY FLEFGTSRHSVESLCVPGVVVLQCSTARAYAL LLCHLPHVLWPRYPVRWULGGDLAGGTFARSIQQV VAPLRADGDLYRAHLTVVRQTTTQHRRQTPLL					sequence		
TASGCR/LYDACRKS.RSOVWPPQTKFG/U.  ATVPCRGAGLG.RGAGAVRUEDAGQWLE PDLFNCTSPARELS.LLLDGLEINKTALDTME AKKLAQRIEVTGHTDHYFSQDVRVTARIL AHLLAFESIQQGFGLTATQDAHFNENILLWA GSALLAFETGDLWAALGQRAFGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLIPSQSPRPSPSEVTSSSIENSTTSS VVPPPAPPEPEPGISIIILVYRTI.GGLIPAGFQ AERIGARI.PONPYMANSVAVAVPHGRINILR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGWWTARDCELVIRNOSHARGCSRTGT FGVLMDASPRERLEGDLELAVFTHVVVAVS VAALVLTAAILLSI.RSLKSNNVGHANVAAA LGVAELLFLIGHRTHNOLVCTAVVILLHYFF LSTFAWLFVQGLHI.YRMQVEFRNVGRGAMR FYHALGWGPPAVLLGIALTSSFLILLI.VSA SWLFGILLAVHSILAFHYLHAGLCGLQGLAV LLIFCVI.NADARAAWMPACLGRKAAPEEAR PAPGLGFGAYNNTALFESGLIRITI.GASTVSS VSSARSGRTQDQDSQRGRSYLRDNLVRHIGS AADHTDHSLQAHAGFTDLDVAMFHEDAGA DSGSDSDLSLEERSI.SIPSSESEDNETMRTAFT QRPLCRAAAGSERLLTHFUNDFOCHOLLSYMPA LGECEAAPCALQTWGSERRI.GLDTSKDAAN NNQPDPALTSGDETSI.GRAQRQRKGILKNRL QYFLVPQTRGAPELSWCRAATI.GHRAVPAAS YGRIYAGGGTGSLSQPASSYSSREGLDLLLRR QLSERILEEAPAPVLRFSRPGSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSSCECMDAAPG RLEFKDRGSTLPRRQPFRDYPOAMAGRFGSR DALDI.GAPREWLSTLPSRPGSSCECMDAAPG RLEFKDRGSTLPRRQPFRDYFOCHAAFG RLEFKDRGSTLPRRQPFRDYFOCHAAFG RLEFKDRGSTLPRRQPFRDYFGCSRARYAL LLQHLPVLVWLPRYYFVRDWLLGGLLSGLSVA MQLPGGLAYALLAGLPPVEGLYSSFYPVFIY FLFGTSRBIRVSSLCVPPVDTT TATPSATASVLGFSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYMMERPLLNQEH LEELGRWGSAPRTRQWTRWLCGCRARAYAL LLQHLPVLVWLPRYYFVRDWLLGGLLSGLSVA MQLPGGLAYALLAGLPPVEGLYSSFYPVFIY FLFGTSRBIRVSSLCVPOPVDT FLFGTSRBIRVSSLCVPOPVDT FLFGTSRBIRVSSLCVPQALKGGEFAPSRIQQV YAPLRADGGACVQFQALKGGEFAPSRIQQV YAPLRADGGACVQFQALKGGEFAPSRIQQV YAPLRADGGACVQFQALKGGEFAPSRIQQV YAPLRADGGACVQFQALKGGEFAPSRIQQV YAPLRADGGACVQFQALKGGEFAPSRIQQV YAPLRADGGACVQFQALKGGEFAPSRIQQV YAPLRADGGACVGFQALKGGEF			{ ' '				TNGQCHCKEFHYRPRGSDSCLPCDCYPVGS1
ATYPCPRGALGARGAGAVRLCDEAGGWLE PDLENCTSAPRELSLLOLEINKTALDTME AKKLAQRIREVTGHTDHYPSQDVRVTARIL AHLAFESHQQGFGLTATQDAHFNENLLWA GSALLAFETGDLWAALGQRAPGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGUTYPINML SIDRMEHESSPRGARRYRYHSINLFRGQDAW DPHTHVLLPSQSPRPSPESVLPTSSSENSTTSS VYPPPAPPEPPEGISIIILLVYRTLGGLLPAQFQ AERRGARLPQNPVMNSPVVSVAVHGRNELR GILESHSLEFRLLQTANRSKAICVQWDPPGLA EQHGWVTAARDCELVIRRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAARLISLRSLKANVRGHANVAAA LGVAELLFLIGHRTHNQLVCTAVVILLHYFF LSTFAWLFVGGLHLYRNNYRGHANVAAA LGVAELLFLIGHRTHNQLVCTAVVILLHYFF LSTFAWLFVGGLHLYRNNYGHANVAAA LGVAELLFLIGHRTHNQLVCTAVVILLHYFF LSTFAWLFVGGLHAYAGLERNAVFRGHANVAAA LGVAELLFLIGHRTHNQLVCTAVVILLHYFF LSTFAWLFVGGLHAYAGLISRSFLLLLVSA SWLFGLLAVYHSILAFHYLHAGLCGLQCLAV LLLFCVLNADARAAWMGTMELLA ARTSCSTGGREAKTSALTLRSSFLLLLLVSA SWLFGLLAVYHSILAFHYLHAGLCGLQCLAV LLLFCVLNADARAAWMGTMELA ARTSCSTGGREAKTSASVLRDNIVLYRIGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDSLSEERSLSPSSESDINGTRRGRF QRPLCRAAQSERLLTHFNVDVGGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPPDALTSGDETSIGARQRKGLKINRL QYFLVPQTRGAPELSWCRAATLGHRAYPAAS YGRIYAGGGTGSLSQPASQRKGKLKNRL QYFLVPQTRGAPELSWCRAATLGHRAYPAAS YGRIYAGGGTGSLSQPASQRKGKLKNRL QYFLVPQTRGAPELSWCRAATLGHRAYPAAS YGRIYAGGFTGSLSQPASYSRSEQLDLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEFKDRGSTLPRRQPFRDYPGAMAGRFGISR DALDLGAFREWLSTLPPFRRTMDLDPQPPFLP LSPQRQLSROPLLFSRPLDSLSSSNSRQCLDQ VPSRHPSREALGFLPQLLRAREDSVSCPSHGP STEQLDLSSILASFNSALSSVOSSTLEGHT TATPSATASVLGPSTPRSATSHSSELSPOSSER DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLCGCSARAYAL LLQHLPVLVWLPRTYPVBWLLGDLLSGLSVA MQLPQGLAYALLAGLPPVGGLYSSTYPVFIY FLFGTSRISHVSSLCVPOPVDT FFGTSRSSRSMADDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLFFLAGLTTYL LVSQLRAQGEACVQFQLKLGGEFAPSRQQV YAPLRADGGKPRAHLTVVRQPTTPHFKNQDFT ALHWEHELGLAFTKNRMYTNKFLLIPEGS	1		1			ł	TACCODIT VDA COVEL DECVIVIDATVEGVI
PDLFNCTSPARPELSLLLOGLEINKTALDTME  AKKLAQRLREVTGHTDHYFSQDVRVTARILI  AHLLAFESHOQGFGLTATQDAHFNENLLWA  GSALLAPETGDLWAALGQRAPGGSPGSAGLV  RHLEEYAATLARNMELTYLNPMGLVTPIML  SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW  DPHTHVLLPSQSPRPSSPLTSSSENSTISS  VVPPPAPPEPEGISIIILVYRTLGGLLAQFG  AERRGARLQNDPWMSPVSVAVPHGRNFLR  GILESPISLEFRILLQTANRSKALCVQWDPFGLA  EQHGVWTARDCELVYRROSHARCKCSKTGT  FGVLMDASPRERLEGDLELLAVFTHVVVAVS  VAALVLTAALLSIRSLKSNNVGGHANVAAA  LGVAELLFLIGHRTHNOLVCTAVVILLHYFF  LSTRAWLFVQGLHLYRNQVEFRYNDRGAMR  FYHALGWGPPAVLLGLAVGLDPEGYGRPDF  CWISVHEPLIWSFAGPVVLVIVMNGTMFLA  ARTSCSTGGREAKTSALTSSFLLLLLVA  ARTSCSTGGREAKTSALTSSFLLLLVSA  SWLFGLLAVNHSILAFHYLHAGLCGLQCLAV  LLLFCVINADARAAWMPACLGRKAAPEAR  PAPGLGFGAVNNTALFESGLIRTIGASTVSS  VSSARSGRTQDQDSQRGRSYLRDNLVRHGS  ADHTDHSLQAHAGFTDLDVAMFHRDAGA  DSDSDSDLSLEERSLSPSSESSEDNGRTGGRF  QRPLCRAAQSERLLTHYDVGRNLLSYWPA  LGECEAAPCALQTWGSERRI.GLDTSKDAM  NNQPDPALTSGDETSIGRAGQRKSGLKNRL  QYFLVPQTRGAPELSWCRAATLGHRAVPAAS  YGRIYAGGGTGSLSQPASSYSSREGLDLLRRR  QLSERLEEAPAPVLRFISRPGSQECMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPRRTEDLDPQPPFLP  LSPQRQLSRDPLLPSRPGSGECCMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPRRTEDLDPQPPFLP  LSPQRQLSRPPLLPSRPTSRGSGECMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPSRTSGSGECMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPSRTSRGSGECMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPSRTSBRGSCCMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPSRTSBRGSCCMDAAPG  RLEFKDRGSTLPRRQPFRDYPGAMAGRGGSR  DALDLGARFEWLSTLPPSRTSBRGSSCFRDSEP  DTQALLSATQAMDLRRRDYHMERPLLNQEH  LEELGRWGSAPRTHQWRTWLQCSCRARAYAL  LLQHLPVLYWLPRYPYRWWLLGDLLSGLSVA  MQLPQGLAYALLAGLPPVFGLYSSFYPVFIY  FLFGTSRHIBWSSLCVPPGWALKGGEFAPSHQQV  YAPLRADGGRAYARLTCCLVLLPFLAGLTTYL  LYSQLRAGGEACVQFQALKGGEFAPSHQQV  YAPLRADGGRACKGALKGGEFAPSHQQV  YAPLRADGGRACKGALKGGEFAPSHQQV  YAPLRADGGRACKGALKGGEFAPSHQQP  ALHWEHELELGLATTKNRNYTNKFLLPPSGD		1	ĺ			}	
AKILAQRIREYTGHTDHYFSQDYTARILI ARHLAFESHQQFGLTATQDAIFNENLLWA GSALLAPETGDLWAALQQRAPGGSPGSAGLV RHLEEYAATLARMELTYLNPMGLVTPNIML SIDRMEHPSSRGARRYPRYHSINLFRGQDAW DPHTHVLLFSQSPRSPSEVLPTISSENSTTSS VYPPPAPPEPEGISIIILLVYRTLGGLLPAQFQ AERRGARLPQNPVMNSPVSVAVHGRNFLR GHLESPISLEFRLLQTANSKA/CVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAALILSLRSLKSNVRGHANVAAA LGVAAELLFLLGHRTHNQLVCTAVVLLHYFF LSTFAWLFVQGLHLYRMQVFRNVDRGAMR FYHALGWGYPAVLLGLAVGLDPEGYGPDDF CWISVHEPLIWSFAGFVVLVVIMONTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA LLLFCVLNADARAAWMFQVFRNVDRGAMR FYHALGWGYPAVLLGLAVGLDPEGYGPDDF CWISVHEPLIWSFAGFVVLVVIMONTMFLLA ARTSCSTGGREAKKTSALTLRSSFLLLLLVSA LLLFCVLNADARAAWMFGAGFGST SVSARSGRTQDQDSQRGGSVLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHBDAGA DSDSDSDLSLEERSLSPSSESEDNGRTRGRF QRFLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSIGAGRGKGLKNRL QYFLYPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASYSSREGLDLLLRR QUSRERLEEAPAPVLRFLSRPGSGCKDAAPG RLEFKDRGSTLPRRQPRDYPGAMAGRFGSR DALDLGAFREWLSTLIPFRRTTBLDPQFPFLP LSPQRQLSRDPLLFSRPLDSLSSSNSEQLDQ VPSRHPSREALGFLPQLILRAREDSVSCPSHGP STEQLDLSSILASFNSALSSVOSSTLIGHTT TATPSATASVLGFSTPRSATSHSISELSPDSER DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPTTRQWRTWLCGSARAYTAL LLQHLPVLVWLPRTYPVRWLLGGLLSGLSVA MQLPQGLAYALLAGLPPVEGLYSSFTYPTF FLFGTSRINSVSSLCVPPDVDT FFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSRWALDCLULFFLAGLTTYL LVSQLRAGGEACVQFQALKGGEFAPSTQQV YAPLRADGGKPRAHLTVVRQTTTPHFKNQFP ALHWEHELGLAFTKNRMYTNKFLLIPFESG	i		[			Í	DOI ENCTSDARDER SELL DOI ELNKTALDTME
AHLLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAGQRAPGGSSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHFSSPRGARRYPRYHRIAFRQQDAW DPHTHVLIPSQSPRPSPSEVLFTSSSENSTTSS VVPPPAPPEPEPGISIIILLYYRTLGGLLPAQFQ ARRGARLPQNPVMAYPYSVAVFHGNIPLR GILSPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCCSRTGT FOVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGHANVAAA LGVAELLFLLGHRITHNOLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGIDPGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSLLFHYHAGLCGLQLAV LLLFCVLNADARAAWMPACLGGLAAPEEAR PARGLGPGAYNNTALFESGLBRITLGASTVSS VSSARSGRTQDQDSQRGSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DDSDSDDSLSLEERESJFSSESEDNGRTRGRF QRFLCRAAQSERLLTHPKDVDONDLLSYWPA LGECRAAPCALQTWGSERRLGLDTSKDAAN NNQPDFALTSGDETSLGRAQRQRKGULKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGNYAGGGTGSLSQPASRYSSRBQLDLLLRR QLSRERLEAPAPVLRFLSRPGSQECMDAAPG RLEPKDRGSTLPRRGOPPGDVPGMAGRGFGSR DALDLGAPREWLSTLPPPRTTRDLDPOPPFLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHFSREALGPLPQLLRAREDSVSGPSHGP STEQLDLLSSTLASRNSSALSSVQSSSTLGPHT TATFSATASVLGGFTRSATSHSISELSPDSEFR DTQALLSATQAMDLRRDYHMGRPLLNQEH LEELGRWGSAPRTHQWTRWLCGCRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYVFIY FLGTSRTHISVESLCVCTGVTSTFYFFIT FLGTSRTHISVESLCVCTGVTSTFYFFIT FLGTSRTSISSSALSGLAGLTSVA IMQLPQGLAYALLAGLPPVFGLYSSFYVFIY FLFTSRTHISVESLCVCTGVTTYTCHFKNQPY ALFWEHELGLAFTKORMYTNKFLLIPESGD				1			AKKI AORI REVTGHTDHYESODVRVTARIL
GSALLAPETGDLWAALGQRAPGGSPGSAGUV RRILEEYAATLARNMELTYLNPMGUVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVILPSQSPRPSPSEVLFTSSIENSTTSS VVPPPAPPEPGTGISIILLYYRTLGGLIZAQFQ AERRGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVTTHVVVAVS VAALVLTAAILLSLRSLKSNVRGIHANVAAA LGVAELLFLLGHRTHNOLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVPSRNVDRGAMR FYHALGWGVPAVLIGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLGCIQGLAV LLLFCVLNADARAAWHPACGLCRKAAPEAR PAPGLGPGAYNNTALFEESGLINTLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AAPHTDHSLQAHAGTDLDVAMFHRDAGA DESDSDDLSLEERSLSPSSESEDNGRTRGRF QRPLCRAAQSERLITHPKOVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRGRGILKNRL QYLVPOTTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRPLLESWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRPLLESWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRPLLESRFRDSCSCEMDAAPG RLEPKDRGSTLPRRQPFRDYFGAMAGRGSR DALDLGAPREWLSTLPPPRRTTDLDPQPPPLP LSPQRQLSRDPLLFSRFLDSLSSSNSSREQLDL VPSRHPSREALGPLPQLLRAREDSVSCPSHGP STEQLDLSSLASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEFR DTQALLSATQAMDLRRRDYHMEPPLLQGEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVEDWLLGGLLSGLSVA MQLPQGLAYALLAGLPPVFGLYSSFYPFIF FLGTSRISVSELCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT FLGTSRTIISVESLCVPGPVDT ALPWLAGGTAVLLICHTTYLL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKRRAHLTVVRGTTTYHKNQTTY	ļ		1	}			
RELEEYAATLARMELTYLIPMGLVTPNIML SIDEMEHPSSPRGARRYPKINSLERGQDAW DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS VVPPPAPPEPEPGISIIILLVYRITIGGLLPAQFQ AERRGARLPQNPVMNSPVVSVAVFHGNISLER GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHQVTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTTHVVVAVS VAALVLTAAILLSLRSLKSNVRGHANVAAA LCVAELLFLLGHRITHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRIVDRGAMR FYHALGWGPAVLUGLAVGIDPEGYGPPDF CWISVHEPLIWSPAGPVVLVVMNGTMFLLA ARTSCSTQRAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DSDSDSDLSLEERESJFSSESEDNGRTRGRF QRFLCRAAQSERLLTHPKDVDONDLLSYWPA LGECRAAQSERLLTHPKDVDONDLLSYWPA LGECRAAQSERLLTHPKDVDONDLLSYWPA LGECRAAQSERLLTHPKDVDONDLLSYWPA LGECRAAQSERLLTHPKDVDONDLLSYWPA LGECRAAQSERLLTHPKDVDONDLLSYWPA ROPPOPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGNIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEAPAPVLRFLSSPGSGCCMDAAPG RLEPKDRGSTLPRRGDSGCCMDAAPG RLEPKDRGSTLPRRGDLSSRSNSSEQLDQ VPSRHPSREALGPLPQLLRAREDSVSQPSHGP STEQLDLLSSILASFNSSALSSVQSSSTELGHTT TATPSATASVLGPSPTDLSFRSSNSSEQLDQ VPSRHPSREALGPLPQLLRAREDSVSCPSHGP STEQLDLLSAILASRYSTSHSISELSPDSEFR DTQALLSATQAMDLRRRDYHMGRPLLNQEH LEELGRWGSAPRTHQRTWHUCCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYVFIY FLFGTSRHISVESLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT FLFGTSRHSVSSLCVPGPVDT ALVERLEGLTTYLLUSQLRAQGGACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRGITTYGHFKNQPT ALWERTHERDELGLTTYLLUSPGLTTYLLUSQLRAQGGACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRGITTYGHFKNQPT ALWERTHCHTCHTTYLLPFSGD						1	GSALLAPETGDLWAALGORAPGGSPGSAGLV
SIDRMEHPSSRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRPSSPEVLPTSSEIRSTTSS VVPPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERRGARLPQNPVMNSPVVSVAVPHGRNFLR GILESPISLEFELLQTANRSKAICVQMDPFGLA EQHGVWTARDCELVIRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAALLSLRSLKSNYRGHANNAAA LGYAELLFLLGHRTHNQL.VCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYMALGWGYPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFILLA ARTSCSTGQREAKKTSALTLRSSFLLLLUSSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVINADARAAWMPACLGRKAPEEAR PAPGLGPGAYNNTALFEESGLRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DSDSDSDLSLEERSLSPSSEEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSVWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDALTSGDETSLGRAQRGRKGILKNRL QYPLVPQTTGAPPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEAPPAPLRPLSRPGSGCEMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPOPPPLP LSPQRGLSRDPLIPLSRPGSGCEMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPOPPPLP LSPQRGLSRDPLIPLSSRSSNSREQLDU VPSRHPSREALGPLPQLARAEDSVSGPSHGP STEQLDLSSILASFNSSALSSVQSSSTFLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDVHMERPLLNQEH LEELGRWGSAPRTHOWNTWL-QCSRARYALL LLQHLPVLWLPRYPVRDWLLGDLLSSLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLIPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNOPP ALHWEHLGLLAFTKNRMNYTNKFLLIPESGG			1			Ì	
DPHTHVILPSQSPRPSPSEVITSSISIENSTISS VVPPPAPPEPPGISIIILVYRTIGGLPAQFQ AERRGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFILLQTANRSKALCVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVITAAILLSLRSLKSNVRGHANVAAA LGVAELLFLLGHRTHNQLVCTAVVLLHYFF LSTFAWLFVQGLHLYRMQFPRNVDRGAMR FYHALGWGYPAVLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGGLAVNHSILAFHYLHGGLGQLGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLHITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DSDSDSDLSLEERSLSJSSSSESEDNGRTRGRF QRFLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERLCLDTSKDAAN NNQPPALTSGDETSLGRAQRQRKGLKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGTGSLSQPASRYSSREQLDLLRR QLSRERLERAPAPVLRFLSRPGSQECMDAAPG RLEFKDRGSTLPRQPPRDYPGANAGRFGSR DALDLGAFREWLSTLPPPRRTNDLDPOPPPL LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLAREDSVSGPSHGP STEQLDLSSLLASFNSALSSVQSSSTLGPHT TATPSATASVLGPSTPSATSISISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLCGSRARAYAL LLQHLPVLVWLPRYPVRDVLGGDLSSVA IMQLPQGLAYALLAGLPPVFGLYSSTPPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSSSMAEDLGLSFGTEASVEMLPEHG SCRPKARSSSARWALTCCLVLLIPFLAGLTTYL LVSQLRAQGEACVQFQALKGQFEAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHLGLLAFTKNRMNYTNKFLLIPESGD	ì		1	1		1	SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW
AERRGARI.PQNPVMNSPVVSVAVPHGRNFIR GILESPISLEFRILQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILSIKSLKSNVRGHANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYKMQUERNVDRGAMR FYHALGWGYPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSPAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLVSA SWLFGILAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEAR PAPGLGPGA YNNTALFEESGLIRITLGASTVSS VSSARSGRTQQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSLSLEERSLSFSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSVWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NQPPPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLAR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEFKDRGSTLPRQPFRDYPGAMAGRGSS DALDLGAPREWLSTLPPPRRTRDLDPOPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEGLDLLSSLRSNSSALESVOSSSTPLGPHT TATPSATASVLGPSTPRSATSHISISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPYVBWLLGDLLSGLSVA IMQLPQGLAYALGGLPSFGTASVEMLPEHG SCRRKARSSSRAWAEDLGJSFGETASVEMLPEHG SCRRKARSSSRAWAEDLGJSFGETASVEMLPEHG SCRRKARSSSRAWAEDLGJSFGETASVEMLPEHG SCRRKARSSSRAWAEDLGJSFGETASVEMLPEHG SCRRKARSSSRAWAEDLGJSFGETASVEMLPEHG SCRRKARSSSRAWAEDLGJEFGETASVEMLPEHG SCRRKARSSSRAWAEDLGJLTYFLLVSGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGKPRAHLTVVRQTPTQHFKNQFP		1				ļ	DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS
GILESPISLEFRLLQTANRSKAICVQWDPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGHIANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSAGPVVLVIVMNGTMFLIA ARTSCSTQQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGL.CGLQGLAV LLLFCVLNADARAWMPACLGRKAAPEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQPQDSOGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMPHRDAGA DSDSDSDLSLEEERSLSFSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASSYSSREQLDLLLRR QLSRERLEAPAPVLRPLSRPGSOECMDAAPG RLEPKDRGSTLPRQPPRDYPGAMAGRTSGS DALDLGAPREWLSTLPPPRRTRDLDPQPPLP LSPQRQLSRDPLLPSRFLDSLSSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHQ VPSRHPSREALGPLPQLLRAREDSVSGPSHQ TATPSATASVLGPSTPRASTHSISELSPDSEFR DTQALLSATOAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPYVBWLLGDLLSGLSVA MQLPPGGLAYALLAGLPPVYRGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRRKARSSSRAWADLTCCLVLLPPLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAFSHQQV YAPLRADGKKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1	ļ					
EQHGWTARDCELVHRNGSHARČRCSRTGT FGVLMDASPRELEGDLELLAVFTHVVVAVS VAALVLTAAILSLRSLKSNVRGHANVAAA LGVAELLFILIGHRTHNQLVCTAVVILLHYFF LSTFAWLFVQELHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVLVIVMNGTMFLLA ARTSCSTGGREAKKTSALTIRSSFLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVINADARAAWMPACLGRKAAPEAR PAPGLGPGAYNNTALFEESGLRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSSDSLSLEEERSLSPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGRDLLSYWPA LGECEAAPCALQTWGSERRJGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSRGQLDLLLRR QLSRERLEEAPAPVLRPLSSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSRGQLDLLLRR QLSRERLEEAPAPVLRPLSSRPSSSSEDNGRFRGF RLEPKDRGSTLPRRQPRDVPFOAMAGRFGS BALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSQPSHGP STEQLDLLSSLASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAVALLAGLPPVGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRNYTNKFLLPPSGD	1		}				1
FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSIKSNVRGIHANVAAA LGVAELIFLIGHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLUSA SWLFGLLAVNHSILAFHYLHAGL.GGLQGLAV LLLFCVINADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLBRITLGASTVSS VSSARSGRTODDSORGRSYLRDNVLVHGHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRGRKGILKNRL QVPLVPQTRGAPELSWCRAATLGFRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRRLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPFRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRTRADDPOPPPLP LSPGRQLSRDPLLPSRDDSLSRSNSREQLDQ VPSRHPSREAGLPPQFLLRAREDSVGSPSHGP STEQLDLSSILASFNSSALSSVOSSSTPLGPHT TATPSATASVLGPTSPATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLCGSRARAYAL LLQHLPVLVWLPPYPVRDWLLGDLSGLSVA MQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMYTNKFLLPPSGD		1					
VAALVLTAAIILSLRSLKSNVRGIHANVAAA LGVAELIFLLGIHRTHNQLVCTAVVILLHYFF LSTRAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLAA ARTSCSTGQREAKKTSALTLRSSFLLLLUSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVIMADARAAWMPACLGRKAAPEAR PAPGLGPGA YNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSLSLEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSCEMDAAPG RLEPKDRGSTLPRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSKDPLSTLPPPRRTRDLDPQPPPLP LSPQRQLSKDPLSTLPPSRPDLDSLSRSSNSTSEQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDLSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHGWRTWLQGCSRARAYAL LLQHLPVLVWLPRYPYRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT FLFGTSRHISVESLCVPGPVDT STEGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSAWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRNYTNKFLLPPSGD			1				
LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHILYRMQVEPRNVDRGAMR FYHALGWGYPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSPAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDCNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPFRDVPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDLSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQCLAYALLAGLPPVFGLYSSFYPVFIV FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP	}		j	]		1	
LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGYPAVLIGLAVGLDPEGYGNPDF CWISVIEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEE\$GLIRITLGASTVSS VSSARSGRTQDQDGQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DSDSDSDLSLEERSLSPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDONDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQQRKGLIKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSREALGAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDVPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTTRDLDPQPPPLP LSPQRQLSRDFLLFSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSCPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVOFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	}	1	ļ	-			
FYHALGWGVPAVILGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTQOREAKKTSALTLRSSFLLILLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS ADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDDSDLSLEEERSLSPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DADLLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRCHALGSLASSANSKEQLDQ VPSRCHAGAGGTGSLSQPASRYSSREQLDQ VPSRCHAGAGGTGSLSGPSTFLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSLATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGBCACVQFQALKGQEFAPSHQQV YAPLRADGBCACVQFQALKGQEFAPSHQQV YAPLRADGBCACVAGTARMNYNKFLLIPESGD		ļ	1 :	1			1
CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTOQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGFGAYNNTALFEESGLIRTLGASTVSS VSSARSGTQDQDSQRGSYLRDNVLVRHGS AADHTDHSLQAHAGFTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTGRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLTPPRRTPDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSAWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGGDRRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLPESGD			1			i ,	
ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGL.CGLQGLALAV LLLFCVLNADARAAWMPACLGRKAAPEAR PAPGLGPGAYNNTALFEESGLRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLFFKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSSSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTFLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLITVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRNNYTNKFLLIPESGD			1	]	}	j	
SWLFGLLAVNHSILAFHYLHAGI.CGLQGLAV LLLFCVINADARAAWMPACLGRKAAPEEAR PAPGGPGAYNNTALFEE\$GLBITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDDSDLSLEERSLSBSESSESDRGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLPQPPPL LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSCPSHGP STEQLDLSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPPSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD		-		İ	l	Į.	
LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLIRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  806 EFGTSRSSSSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1				SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV
VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLFSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD						Ì	LLLFCVLNADARAAWMPACLGRKAAPEEAR
AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDSLSLEEERSLSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQKKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTOHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1	į.		1	PAPGLGPGAYNNTALFEESGLIRITLGASTVSS
DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTIDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCLVLLIPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1		1	1	ł	İ	
QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDULLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1				
LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD		1				1	
NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			ĺ		1		
QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD				}	ļ		
YGRIYAGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1	•	l		NNQPDPALISGDEISLGRAQRQRAGILANRL
QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1				ļ		
RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1		1	ļ		}	
DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD				1		İ	
LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD				1			
VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD				1	1		
TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1			j	
DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	ŧ			i		ļ	STEQLDILSSILASFNSSALSSVQSSSTPLGPHT
LEELGRWGSAPRTHQWRTWLQCSRARAYAL   LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA   IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY   FLFGTSRHISVESLCVPGPVDT   FLFGTSRSSRSMAEDLGLSFGETASVEMLPEHG   SCRPKARSSSARWALTCCLVLLPFLAGLTTYL   LVSQLRAQGEACVQFQALKGQEFAPSHQQV   YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP   ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	Į			Į.	l	1	
911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSGLSVALVSQLRAQGEACVQFQALKGQEFAPSHQVVYAPLRADGDKPRAHLTVVRQTPTQHFKNQFPALHWEHELGLAFTKNRMNYTNKFLLIPESGD	[				1		DTQALLSATQAMDLRRRDYHMERPLLNQEH
IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT  911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1			l			LEELGRWGSAPRTHQWRTWLQCSRARAYAL
911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1		1	1	1	1	
911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	1					- [	1 ' '
SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			<del>                                     </del>	1,7000	10.	1006	
LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD	911	2261	A	7890	[ 21	800	
YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1		1		
ALHWEHELGLAFTKNRMNYTNKFLLIPESGD			1				
				1	1		
	}	į.	ļ				YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV
VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ			ļ		1		VITKVTDSYPEPTOLLMGTKSVCEVGSNWFO
PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE	1	1	İ	1			PIYLGAMFSLOEGDKLMVNVSDISLVDYTKE
DKTFFGAFLL	1			1			
912 2262 A 7891 1263 111 ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL	912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL
LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ	1 712			1			LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ
PVPYILKKIFQDREAAATTGVSRDLCYVKELG			-			1	PVPYILKKIFQDREAAATTGVSRDLCYVKELG
			j		1	1	VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL
			1			- }	YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP
ELELALFLVQEPHVWGQITPKPGKMFVLRSV	L					1	ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

						Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence	<b>!</b>		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Jodgaanse	/=possible nucleotide deletion, \=possible
!	{	1		sequence	ļ	nucleotide insertion
			<del> </del>	Boquoneo		PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL
	1	1	1			EILVKEDRDSGVNFQPEDTCARLRCSLHASLL
	ł	ł				VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH
	i				1	RHOLFINFRDLGWHKWIIAPKGFMANYCHGE
	}	1		į.		CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV
	1	ļ.			ł	CIPTKLSPISMLYQDNNDNVILRHYEDMVVD
	ĺ	1		1	i	ECGCG
913	2263	A	7892	15	849	ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL
		1				YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE
	1	1	1			LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ
		1		}	1	GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS
			1	ļ		VGATCCYLLSSIFGKQLVVSYFPDKVALLQR
		1		1 .	1	KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI
						LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS
		1			İ	LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ
	İ		<u> </u>			KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN
						RLSNITRPFFSKKKKILV
915	2265	A	7909	3.	641	HASGPGGLLRRRRGSGANMPVARSWVCRKT
	,	1	1		1	YVTPRRPFEKSRLDQELKLIGEYGLRNKREV
	i		-	ì		WRVKFTLAKIRKAARELLTLDEKDPRRLFEG
			-	1	1	NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR
		ì				EOVVNILFFTVRLDSQKHIDFSLCFPIGVANPS
						HVKRKNASKGQGGAGARDDEEEE
		-	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC
916	2266	A	/914	3	307	OVLILKHTHASLSLPSCQECFPSSIPSASHMVS
· ·	1	1	1		1	HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS
l		}	j	}	ļ	SILPTGDSWGMLACLCTVLWHLPAVPALNRT
1	1	1		Ì		GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN
						YLGPPFNEPDFNPPRLGAETLPRATVDLEVW
			}			RSLNDKLRLTONYEAYSHLLCYLRGLNRQAA
	1		İ	1		TAELRRSLAHFCTSLQGLLGSIAGVMAALGY
1		1	1			PLPOPLPGTEPTWTPGPAHSDFLQKMDDFWL
İ						LKELQTWLWRSAKDFNRLKKKMQPPAAAVT
		j				LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL
1		1				HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ
1		l			-	LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED
1		1				THHQIATLTVLVAPENPVVEVREQAVEGGEV
}	-	1	1		1	ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ
		1				ENGKVWSVASTVRFRVDRKDDGGIIICEAQN
İ				1	{	QALPSGHSKQTQYVLDVQYSPTARIHASQAV
1						VREGDTLVLTCAVTGNPRPNQIRWNRGNESL
1		1		1		PERAEAVGETLTLPGLVSADNGTYTCEASNK
1			1	1		HGHARALYVLVVYGESRLRPTEGGGGAPDP
1		1		1	1	GAVVEAQTSVPYAIVGGILALLVFLIICVLVG
1		1		1		MVWCSVRQKGSYLTHEASGLDEQGEAREAF
						LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK
1					]	ESPRWRSALLLLFLAGVYGNGALAEHSENVH
1		1	1	1		ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK
1.				Į.		INCSWFIRANPGEIITISFQDFDIQGSRRCNLD
1			1	1		WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR
1	1		1	1	į	FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR
1						
	İ			ł	i	CGNGKCIPEAWKCNNMDECGDRSDEEICAKE
						ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ing to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \(\)=possible nucleotide insertion
						VILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV NAARGFNATYQVDGFCLPWEIPCGGNWGCY TEQQRCDGYWHCPNGRDETNCTMCQKEEFP CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF CQPGNFHCKNNRCVFESWVCDSQDDCGDGS DEENCPVIVPTRVITAAVIGSLICGLLLVIALG CTCKLYSLRMFERRSFETQLSRVEAELLREA PPSYGQLIAQGLIPPVEDFPVCSFNQASVLENL RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA RSRHSGSLALVSADGDEVVPSQSTSREPERNH THRSLFSVESDDTDTENERRDMAGASGGVAA PLPQKVPPTTAVEATVGACASSSTQSTRGGH ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL ASDQGQGLRQPYNATNPGVRPSNRDGPCERC GIVHTAQIPDTCLEVTLKNETSDDEALLLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
920	2270	A	7953	47	572	GGWNDVACHTTMYFMCEFDKKNM GGRASWPEOAKEPRREGHTDKOOTEDVLAA
720	2270		,,,,,,			GLRCLPHLPAICARRMSPAFRAMDVEPRAKG VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS FSSSWPSAQHLTPSVFNPW
921	2271	A	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF MMVYSFRALSFKESTWATFQHGGEATKSRSL SSTQ
922	2272	A	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN WELVKPN
923	2273	A	7981		3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL LLPLLLLPAGCRALEETLMDTKWYTSELAWT SHPESGWEEVSGYDEAMNPIRTYQVCNVRES SQNNWLRTGFIWRRDVQRVYVELKFTVRDC NSIPNIPGSCKETFNLFYYEADSDVASASSPFW MENPYVKVDTIAPDESFSRLDAGRVNTKVRS FGPLSKAGFYLAFQDQGACMSLISVRAPYKK CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN AVEVSVPLKLYCNGDGEWMVPVGACTCATG HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP NSRTTSPAASICTCHNNFYRADSDSADSACTT VPSPPRGVISNVNETSLILEWSEPRDLGVRDD LLYNVICKKCHGAGGASACSRCDDNVEFVPR QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG IASTVTSQMNSVQLDGLRPDARYVVQVRART VAGYGQYSRPAEFETTSERGSGAQQLQEQLP LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL KQPGRREVFVAIKTLKVGYTERQRRDFLSEA SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM KYLSEMNYVHRDLAARNILVNSNLVCKVSDF GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI AYRKFTSASDVWSYGIVMWEVMSYGERPY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  WDMSNQDVINAVEQDYRLPPPMDCPTALHQ LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD
924	2274	A	7985	1	503	WLDAIKMGRYKESFVSAGFASFDLVAQMTA EDLLRIGVTLAGHQKKILSSIQDMRLQMNQT LPVQV FRPRTKKATAMYLEHYLDSIENLPCELQRNF
924	22/4		17905	•		QLMRELDQRTEDKKABIDILAAEYISTVKTLS PDQRVERLQKIQNAYSKCKEYSDDKVQLAM QTYEMVDKHIRRLDADLARFEADLKDKMEG SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS EEDTPKKKKHKGG
925	2275	A	7994	447	589	LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR SPIRCYCOHWPHCVHC
926	2276	A	7996	925	582	GPCKVCCITLAIMLQCHSFYRKDVQVEHPKS LNPKYSQIENFLSADMALKRKCLLSISDLDFW IWDAQPVGIMQTLQNLKKIPNPGCFWSQAFQI RDTQPILPLGGRYYITIRQ
927	2277	A	7998	2	353	RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV CLLLVTLALCCYQANAEFCPALVSELLDFFFI SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLVKILKKCSV
928	2278	A	8004	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG ARGLRATYHRLLDKVELMLPEKLRPLYNHPA GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVGAAGASQLFRIWRYNQELKAKAHK
929	2279	A	8007	2	1016	EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNITVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE
930	2280	A	8008	3	1679	NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL, YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP
931	2281	A	8009	861	300	AAGAVVSAMPKAKGKTRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

				** ** . 1		Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Aianine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ł	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	ŀ	USSN	location	corresponding	
seq-	иепсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
uence	1	1	914	ng to first	acid residue	
	İ		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	ł	l .	peptide	İ	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			1			PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
	1		1	}	)	LIDYVRYMVENHGEDYKAMARDEKNYYQD
		<b>[</b>	İ			TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
		İ				MEVE
932	2282	Α	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
	}	1	1	1	1	DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
	1	ł	1			QGRGQIPIPCPWPPPPPPPPPPGSPGPGCRQFHQ
	Í	1	į.			SLEAKARHPASVREMRGKVKMRRALRRAPA
						STRASSRQPNPK
933	2283	A	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
-		1	1			ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
	1	1	1	J		NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
		ļ	1			RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
!		1	1	ì		PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
	Į	1		1		NASQLITQRAQVSLLIRRELTERAKDFSLILDD
1			1			VAITELSFSREYTAAVEAKQVAQQEAQRAQF
1				ł		LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
		}	ŀ		1	SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
		1		1	-	DNLVLNLQDESFTRGSDSLIKGKK
934	2284	A	8023	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
754	220.	1	0020			RLRKFRELHLMRNEARKLNHQEVVEEDKRL
ļ	j	1	1	1	1	KLPANWEAKKARLEWELKEEEKKKECAARG
	1		1			EDYEKVKLLEISAEDAERWERKKKRKNPDLG
	ì	ł	ŀ			FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
	[	1	1	1		KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
1	1	1		1	ſ	KOIEKRDKYSRRRPYNDDADIDYINERNAKF
			İ			NKKAERFYGKYTAEIKQNLERGTAV
935	2285	A	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
933	2203	A	8027	39	310	OGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
Į		İ	1			SQHSSPAPMYSQTFHILVLG
036	2286	+	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
936	2286	A	0032	1	039	FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
	1	ł		l	}	IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
İ	1	1		İ		NEKSFONIRNWIRNIEEHASADVEKMILGNKC
	1	Ì			ĺ	DVNDKRQVSKERGEKLALDYGIKFMETSAK
1	-		1	1	}	ANINVENAFFTLARDIKAKMDKKLEGNSPQG
	ł	İ	ı	1	[	SNQGVKITPDQQKRSSFFRCVLL
		<del></del>	1	200	211	
937	2287	A	8039	393	311-	EETIHSENSYILEKYIPISANLTLTIA LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
938	2288	Α	8052	675	.1334	
	i	1		1	1	PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
1		1		1	1	ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
	ĺ	-		1	1	LREYQTRQDQCIYNTTYLNVQRENGTISRYV
	ł	į.				GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
1	I	1	1		1	GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
L	<u> </u>	1 -				VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	A	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
1					i	AEQLKWSAELARLGESIMDGKQGGMDGSKP
1	1	1	1	1	1	AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
1		1	1			IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
1			1	1		CLGGHLSCVKILLKHGAQVNGVTADWHTPL
1	.	-		1		FNACVSGSWDCVNLLLQHGASVQPESDLASP
		1		1	1	IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
1	1	1	1	1		PLYLACENOORACVKKLLESGADVNQGKGQ
				1	1 .	DSPLHAVARTASEELACLLMDFGADTQAKN
}		1	Į	1	}	AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
	1	i	1	1	ł	CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
		1		1		L
940	2290	HA-	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
770	2250	( ' '	5555	1 -	1	ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI
L						

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
						ANSVVVWVNIQAKTTGYDTHCYILNLAIADL WVVLTIPVWVVSLVQHNQWPMGELTCKVTH LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS RKKMVRRVVCILVWLLAFCVSLPDTYYLKT VTSASNNETYCRSFYPEHSIKEWLIGMELVSV VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI PFTCRLEHALFTALHVTQCLSLVHCCVNPVL YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA SRVSETEYSALEQSTK
941	2291	A	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLIATANQGSEATYESGF VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA VFEPSEAEISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSGVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPVTQIVSAEAWGRADCGFTS ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG
943	2293	A	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF
944	2294	A	8073	1	797	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM
945	2295	A .	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCORLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE
946	2296	A	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR
947	2297	A	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
948	2298	В	8093	3905	846	PCPPAIMYQSSNKC  MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEBASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGEGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKGFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTTTKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVTYLSCTSELDRLLSALNSGWKTTY QVDLPHTTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKIAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKRHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVYAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD NPMSTPKDSNPKA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Į į	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ			1	peptide	•	/=possible nucleotide deletion, \=possible
Ì		ļ		sequence		nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE
1		ļ		ļ		PPADEAARAGEGFRYIKPVPGLLLREYLYGG
i	!			(	i	GRDEEPSGAAPEGGATPTAAPETPAPPTRETC
			1			YFLNATILFLFRELRDTALTRRWVTKKIKVEF
			1	<u> </u>		EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI
	i	}	ì	}		RLVRPVVPSATGEPDGPEGEALPAACPEELAF
	ł	1				EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS
ì	ļ			ļ		RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV
Ì	}	1	1			RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY
	1.		1		ĺ	KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE
1	1	}				GRLKVTLLECSRLLIFGSYDREANVHCTLELS
1	1		1			SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE
1				}		AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ
	1	1				CPG
951	2301	Α	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR
		1				NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR
İ	1	1	}	Ì		SLSSEKLALLKQIQEAYGKCKEFGDDKVQL
ł	1	1	}	ł		AMQTYEMVDKHIRRLDTDLARFEADLKEKQI
ł	1					ESSDYDSSSSKGKKKGRTQKEKKAARARSKG
ŀ		1	1			KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF
İ	1					GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC
[	1	1	1			
		<del> </del>	10110	500	291	PRCSQERKKK PSVASLARRFSGRALWPPSHSVPGNRALCPRL
952	2302	A	8112	595	291	LHGTTLPGGNQRELARQKNMKKQSDSVKGK
1	1	1		i .		RRDDGLSAAARKQRDSTPRDSEIMQQKQKK
		1		1	1	ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG
755	2505	1 ~	0110	1.	1005	LETNILKMTTPNKTPPGADPKQLERTGTVREI
		1				GSOAVWSLSSCKPGFGVDQLRDDNLETYWQ
						SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE
		1				SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW
•		1				IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD
						THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS
		1		i		IR.
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA
				1		ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE
1	1		١.	1		AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV
1	1	1	1		†	PVGQRRAWCWCMCFGLAFMLAGVILGGAY
1	}	1		1		LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD
1	1		1	1	1	APAALYQTIEENIKIFEEEEVEFISVPVPEFADS
			1			DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT
1	1 .			1		SIVMPPRNLLELLINIKAGTYLPQSYLIHEHMV
		1	1	1	1	ITDRIENIDHLGFFTYRLCHDKETYKLQRRETI
L				<u></u>		KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV
1			ļ		1	KKKIKEVTEEVANKVSCAMTDEICRLSVLVD
		1	1	1	1	EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL
						ADRCTDEVNALVLQTQQEIENLKPLLPAGIQ
	1	1		1	1	DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV
1	1	1	ļ	}	}	FRESLGWSSLVHRFLGPRNAQRVLLGLSEPIF
	-	1	1		1	QLPRSLASTPTAPTTPATPDNASQEELMITLVT
			1	1	1	GLASVTSRTSMGIIVGGVIWKTIGWKLLSVS
				1		LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL
1	1	1	1	1		CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS
İ	1			1		KLLRNKAVQLENELENFTKQFLPSSNEES
055	2206	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH
956	2306	\ ^	0121	1034	1/3	LFLLTAGPALGWNDPDRMLLRDVKALTLHY
L						

						Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	i	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	ŀ	USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino	
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i	ł	ŀ	}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	ļ		peptide	1	/=possible nucleotide deletion, \=possible
	j		}	sequence		nucleotide insertion
					[	DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI
			1	· ·	1	QCQNKGWDGYDVQWECKTDLDIAYKFGKT
		1		1	]	VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL
	1		1		1	GLQKLKESGKQHGFASFSDYYYKWSSADSC
	,	İ		Į.		NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP
		1	l		1	YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ
	}	1			ļ	NTGHGATSGFGSAFTGQQGYENSGPGFWTGL
	1			ł		GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY
	1	ĺ		į		PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA
			1	1		SGYGGTRRR
		<u> </u>		1100	500	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL
957	2307	Α	8159	1492	528	
	1	1				VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ
I .	1	1		1		DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV
	<b>j</b>	1	1	ļ		PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF
	1					QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA
	1		1		Ì	HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG
ļ				1		TLVGLGLLAGLVLLYHRRGKALEEPANDIKE
	(		İ	1		DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL
	1	1	1	}	1	RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG
[	1	1		l	ĺ	AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS
ŀ	1		ł	i		OAGSLV
050	2308	A	8161	2340	1192	ELARRPKOOSSEKSRNMIRNWLTIFILFPLKLV
958	2300	^	8101	2540	1172	EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP
	į		1		1	LNATLVITFEITFRSKNITILELPDEVVVPPGVT
l	1	1	1	1		NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR
1	1		1	Ì	1	M22LA 126MAGGT A LEUGHURAGA KIK
i	1	1	1			FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM
1		1	1			NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL
	1	1	1	ļ		LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH
1		1				AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL
	1	)	1		ļ	AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL
		1				AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL
	Ì	Į.	1	1		DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK
	1			ļ		FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE
737	20	1 **	0.05	1	1	MOPPGPPPAYAPTNGDFTFVSSADAEDLSGSI
į.	1		j	1	ì	ASPDVKLNLGGDFIKESTATTFLRQRGYGWL
ł	1	ł	}	1	1	LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV
	ì	1	Ì		İ	LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS
1		1			1	MISLYGOFRVVSWIITIWIFGSLTIFLLARVLG
}	· ·	1			}	GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF
1	1	1		1	1	EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK
1		1		1	ł	KPLLIYPIFLLYIYFLSLYTGV
				<del></del>	1	
960	2310	A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN
	1	1	1	1	}	LYSQLNALQFTVDERSILWLNQFLLDLKQSL
1		1		1		NQFMAVYKLNDNSKSDEHVDVRVDGLMLK
		1	1			FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH
1				1	1	CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS
1	1	1	1	{		CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP
1	[			1		QLNKNTLKTSAATDVWAVYFSQFWIDYEGM
				}	1	KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP
}	j	1	1	}	1	QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY
1	1	Į.		1		YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI
1	}			(		HLLVHVHKHVSMQINHYQYLLLLFLHESLILL
1			1	1		SENLRKDVEAVTGSPASQTSICIGILLRSAELA
	-	1		1	1	LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN
1	j		Ì	1	1	PPDEL CONDINO CONDINATION OF STATEMENT OF ST
i		1	1	1	1	GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM
	1	- 1	1	1	1	CARPE OFFINE INTONE I PIECE AND THE OFFINE INTE
		1	1			SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY
						LSDKHLGKISEDESSGLVYKSGSGEIGSETSD

oro m	SEO ID	Mot	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	Met hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lvsine, L=Leucine,
seq-	neuce	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	исисс	ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
пенес			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	soquence	/=possible nucleotide deletion, \=possible
		)		sequence	ļ	nucleotide insertion
		<u> </u>	ļ	sequence		DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP
	1					LCVSYKNMKRSSSOMSLDTISLDSMILEEQLL
		1	1			ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN
		ſ	ļ	ļ	ĺ	AGANLONYGETSPDAISTNSEGAQENHDDLM
		·	ţ			SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP
		l	ļ			DOLGNISLRHYLCNRPVGSDQKAVIHSKSSPE
	1	İ	1	ł.	ļ	ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
						EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT
			1	Į		KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER
			Ì		İ	SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL
		1		1		TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN
	<u> </u>	ì	l	i	i	FPEFSFDFTREQLMEENESLKQELAKAKMAL
	1		1			AEAHLEKDALLHHIKKMTVE
0/1	0231	ļ	8172	1442	682	TAAMSIFTPTNOIRLTNVAVVRMKRAGKRFEI
961	2311	A	81/2	1442	082	1 ***
		1	1			ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG
		1				EVQVSDKERHTQLEQMFRDIATIVADKCVNP
	[	1				ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA
		1				
ļ	1	)				LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL
	1		1	İ		KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI
		<del> </del>	0155	206	502	DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS
ĺ		1		1		VLRRMQKKYWKTKQVFIKATGKKEDEHLVA
ļ		1		1		SDAELDAKLEVFHSVQETCTELLKIEKYQLR
<u> </u>		<u> </u>	0.01	<u> </u>	10015	LNGMKS
963	2313	ķ	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ
i		j	1		}	GMDLVWSAWYGKCVKGKGSLPLSAHGIVV
			1			AWLSRAEWDQVTVYLFCDDHKLQRYALNRI
						TVWRSRSGNELPLAVASTADLIRCKLLDVTG
1		1	1			GLGTDELRLLYGMALVRFVNLISERKTKFAK
l		1	1	Ĭ		VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI
				1		NDCRRGCYFVLDWLQKTYWCRQLENSLRET
			1			WELEEFREGIEEEDQEEDKNIVVDDITEQKPE
)	ļ	1	1	1 .	1	PQDDGKSTESDVKADGDSKGSEEVDSHCKK
	İ		ļ		1	ALSHKELYERARELLVSYEEEQFTVLEKFRYL
		1	i			PKAIKAWNNPSPRVECVLAELKGVTCENREA
			i	1		VLDAFLDDGFLVPTFEQLAALQIEYEENVDL
1.				1		NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE
	1		1	1	1	RMLSELPALGISGIRPTYILRWTVELIVANTKT
	1	1		l	1	GRNARRFSAGQWEARRGWRLFNCSASLDWP
				1		RMVESCLGSPCWASPQLLRIIFKAMGQGLPD
]	1	]	1	1	1	EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK
1	1			Į.		SPYTLDSLYWSVKPASSSFGSEAKAQQCEEQ
				1	•	GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE
1	1	1				EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA
		1	1			ENARLLAQKRGALQGSAWQVSSEDVRWDTF
1	1	1	1			PLGRMPGQTEDPAELMLENYDTMYLLDQPV
		1				LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS
<u></u>			<del> </del>	<del> </del>	1000	NFEGLLWSQGQLHGLKTGLQLF
964	2314	Α	8184	6	1393	EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE
1	ł	1			1	GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR
	1		1	1	1	PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP
1		1	1	1	I	NVGKSTFFNVLTNSQASAENFPFCTIDPNESR
	Ì	1			1	VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG
					1	
				,		LVKGAHNGQGLGNAFLSHISACDGIFHLTRA
				,		LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI
						LVKGAHNGQGLĞNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS
						LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS
						LVKGAHNGQGLĞNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS

PCT/US01/03800

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \rightarrowpossible nucleotide insertion  ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV MKYEDFKEEGSENAVKAAGKYRQQGRNYIV EDGDIIFFKFNTPQPKKK RSFSLSFSLLSPSEMMALGAAGATRVFVAMV AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY QPYPCAEDEECGTDEYCASPTRGGDAGVQIC LACRKRKRCMRHAMCCPGNYCKNGICVSS DQNHFRGEIEETITESFGNDHSTLDGYSRTT LSSKMYHTKGQEGSVCLRSSDCASGLCCARH
				1		FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ
	10011	<b> </b>	9005	416	4082	RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW
966	2316	A	8207	416	4002	SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY WVDLERQLLQRVFLNGSRQERVCNIEKNVSG MAINWINEEVIWSNQQEGIITVTDMKGNNSHI LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAK RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHICLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH
967	2317	A	8210	3	601	MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
ectide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of peptide sequence	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFIQTLEDVFRRIFIT YMDNWRQNTTAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATTHENIG AAGFKMSP
969	2319	A	8215		1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR VPGRLLLLLPALCCLPGAARAAAAAAAGGN RAAVAVAVARADEAEAPFAGQNWLKSYGY LLPYDSRASALHSAKALQSAVSTMQQFYGIP VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGQKWRQKHITYSIHNYTPKVGELD TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT KPNICDGNFNTVALFRGEMFVFKDRWFWRL RNRVQEGYPMQIEQFWKGLPARIDAAYER ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG ELGSCLPREGIDTALRWEPVGKTYFFKGERY WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY PRNILRDWMGCNQKEVERRKERRLPQDDVDI MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN YELLEKEVEKRKVALQEAKLKAKGLNPDGTP ALSTLGGFSPASKPSSPREVKAEEKSPISINVK TVKKEPEDRQQASKSYYNGVRKDSKRSRNSR SASRSRSRTRSRSRSHTPRRHYNNRRSRSGTY SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA KKHRHERGHHRDRRERSRSFERSHKSKHHGG SRSGHGRHRR
971	2321	A	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEQESFFEGKNMALFEEEMDSNPMVSSLL NKLANYTNLSQGVVEHEEDEESRREAKAPR MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG VLESFLIVAMCCTCTMLTAISMSAIATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV CLLGNRTLSRRSFDACVKAYGIHNNSATSAL WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA ASGVFLENLWSTYAHAGAFVEKKGVPSVPV AËESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVIVIGSFFSTCGAGLQTLTGAPRLL QAIARDGIVFFLQVFGHGKANGEPTWALLLT VLICETGILIASLDSVAPILSMFFLMCYLFVNL ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL CLALMFICSWYYALSAMLIAGCIYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
		ŀ	09/496			
seq-	uence	ł		correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Į.		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
İ		ł	i I	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i		i i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŧ	1	1	peptide		/=possible nucleotide deletion, \-possible
ļ	1	) :		sequence	]	nucleotide insertion
				7		LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE
	İ			1	1	NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ
1	}		l			
l	i	l	1	i		SAGLGGLKHNTVLMAWPASWKQEDNPFSW
1	1				ļ	KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ
1		l	İ			ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH
1	1	j		]	]	KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF
1					,	LYHLRISAEVEVVEMVENDISAFTYERTLMM
1	J		1	]	j	EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS
	1	1	1			HTAAAARTQAPPTPDKVQMTWTREKLIAEK
1		l	i			1
		l	Ì	1	1	YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV
1	1	1		1	1	KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD
		<u> </u>	L			ENYMEFLEVLTEGLNRVLLVRGGGREVITIYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV
	}	]	1	]	1	RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ
1	1	ŀ	†			VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ
	1	1		[	[	REKANGTTVHVGIHPSKVVITRLKLDKDRKKI
		1		}	1	LERKAKSRQVGKEKGKYKEELIEKMQE
973	2323	A	8237	873	4610	
7/3	2323	^	0231	0/3	4010	GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC
1	ĺ	<b>!</b>	i	1	ĺ	PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG
			1		ŀ	AKAWIMDEEEDAEEEGAGGRQDPSRRSIRLR
1	!	•		ļ		PLPSPSPSAAAGGTESRSSALGAADSEGPARG
	}	1	l		Į.	AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG
		İ	İ	ļ		SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP
	l	1	1		}	GLAAEPERPGASAQPAASPPPPQQPPQPASAS
	Ì	1	i	i		CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG
i	ł	Ì	1	1	ì	
	l		İ		ŀ	QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA
	1	ļ	1	1	•	VEREQERVKSAGFWIIHPYSDFRFYWDLTML
	}	1		]	}	LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD
	1	İ	ŀ	1		TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK
1	)	ļ	}	ļ	ļ	YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK
			}			TARALRIVRFTKILSLLRLLRLSRLIRYIHOWE
1	<b>[</b>	]	l			EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG
	i	İ	l		ļ	CLQFLVPMLQDFPDDCWVSINNMVNNSWGK
ĺ	í	ĺ	l	ĺ'	Í	
	1		}			QYSYALFKAMSHMLCIGYGRQAPVGMSDV
1		Ì	1	1	1	WLTMLSMIVGATCYAMFIGHATALIQSLDSS
1	1		Į	ŀ	1	RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD
	1	l	1			YYEHRYQGKMFDEESILGELSEPLREEINFNC
I	1	l		1	ł	RKLVASMPLEANADPNFVTSMLTKLRFEVFQ
1	1		1		1	PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE
1	1	1	ĺ	İ	l	TKLADGSYFGEICLLTRGRRTASVRADTYCR
	1	l	1		[	LYSLSVDNFNEVLEEYPMMRRAFETVALDRL
ſ	[	l	l	ſ		DRIGKKNSILLHKVQHDLNSGVFNYQENEIIO
	1		1	[		
1	1	1	1			QIVQHDREMAHCAHRVQAAASATPTPTPVIW
Į.	!	į	l	1	1	TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR
1		1	1		<b>f</b>	PPPGSGLGNLGAGQTPRHLKRLQSLIPSALGS
1			[		1	ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS
	l	1	1	1	1	PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA
1	İ		1	ĺ		GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ
1	1	l		1	ļ	AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS
		l		[	1	
1	1		}	<b>J</b>	1	SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP
1	ļ	1	ŀ		ł	PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP
1	1	l		ļ	f	RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR
1	1	l	i	1	{	RGTPPLTPGRLTQDLKLISASQPALPQDGAQT
	1		1	<b> </b>		LRRASPHSSGESMAAFPLFPRAGGGSGGSGSS
		1	i	í		GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP
1		ł	i	1	1	LSLFGARATSSGGPPLTAGPQREPGARPEPVR
		1		1		SKLPSNL
974	2224	<del>                                     </del>	0047	270	469	
7/4	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL
				1		LVPVKDASRICSLTYLLGSHWNNLVVRSPVL
L		L	L	<u></u>		G

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
975	2325	A	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT ILPQELQAWVQEHCPESAEEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS NLTLHYRTHLVDRPYDCKCGKAFGQSSDLLK HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA
	0055	ļ	0065	200	7006	P CAMACWROLDLI I WENT TERRECTOR LIFE
976	2326	A	8257	298	7086	GNMACWPQLRLLLWKNLTFRRRQTCQLLLE VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTTGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTFMENSQEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITTPGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTIILSTHHMDEADVL GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID DRLSDLGISSYGISETTLEEIFLKVAEESGVDA ETSDGTLPARRNRAFGDKQSCLRFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG

PCT/US01/03800

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ĺ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1			[	amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	residue of	sequence	/=possible nucleotide deletion, \=possible
	1	1		peptide		nucleotide insertion
		<u> </u>	<u> </u>	sequence		LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY
		l	(		1	NEQYTFVSNDAPEDTGTLELLNALTKDPGFG
		1	1	İ		TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM
		1	1			DLFQNGNWTMQNPSPACQCSSDKIKKMLPV
		Ì	<b> </b>	ļ	1	CPPGAGGLPPPQRKQNTADILQDLTGRNISDY
}	Ì	ł	ł			LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS
						NTQALPPSQEVNDATKQMKKHLKLAKDSSA
		1		}	1	DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
	ļ		1 '	· ·	i	HAISSFLNVINNAILRANLQKGENPSHYGITAF
j		1	1	}	ł	NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA
l				1		MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI
		1				YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY
İ		1				VSSTNLPVLALLLLLYGWSITPLMYPASFVFK
1		1	ì	į	1	IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ
	İ	1	1	ļ ·		AMADALERFGENRFVSPLSWDLVGRNLFAM
}		1	1			AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
				Ì		DEDEDVRRERQRILDGGGQNDILEIKELTKIY
1		1	1 :			RRKRKPAVDRICVGIPPGECFGLLGVNGAGK
1	1					SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV
1				1		HQNMGYCPQFDAITELLTGREHVEFFALLRG
Į						VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY
	Ì		}	}		SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
						PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
		1	1			EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
	1		ļ		ł	GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK
				}	1	EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH
1			1	1		IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL
				<u> </u>	1567	SLHKNQTVVDVAVLTSFLQDEKVKESYV IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG
977	2327	A	8260	3	1567	YLOGNVNGRLPSLGNKEPPGQEKVQLKRKV
				1		TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM
		1		ĺ	1	SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH
				1		YTYILEVFGPLPAFVRVWVELLURPAATAVIS
1	ļ	1				LAFGRYILEPFFIOCEIPELAIKLITAVGITVVM
Ì	1					VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV
	1	Ì		1		MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
		1				MYAYAGWFYLNFVTEEVENPEKTIPLAICISM
		i				AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT
1	1			1		FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV
1						SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV
İ						LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA
1						VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII
		1	1			WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
	-		0001	2	2165	RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL
978	2328	A	8261	1 4	2105	RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP
		1				LADAASMSGVRAVRISIESACEKQVHEVGLD
	1	Ī		ł		GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE
	}		}	1	ł	EEAAGTEGDAQEWPGAGSSADQDDEEGVVK
				ŀ		FOPSLWPWDSVRNNLRSALTEMCVLYDVLSI
				1	- (	VRDKKFMTLDPVSQDALPPKQNPQTLQLISK
1		1		1	Ì	KKSLAGAAQILLKGAERLTKSVTENQENKLQ
1		1		1	1	RDFNSELLRLRQHWKLRKVGDKILGDLSYRS
		l	Į.	1		AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL
		,				DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
						LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI
		1	1	1		FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ
1	{	1		1	1	LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE
		1		1		HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASATTVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289	2	1053	FVWNPRGGRKRRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLFVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL
982	2332	A	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFTYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY KAYLODLVEGMDFQGPGBS
984	2334	A	8321	1	1243	ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRLAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	A	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	A	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion  LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH OFKLCIMRRSKGRAEKS
991	2341	A	8369	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE DATLENGCLWFIPGSHTSGVSRRMVRAPVGS
						APGTSPLGSEPARDNSLFVPTPVQRGALVLIH GEVVHKSKQNLSDRSRQAYTFHLMEASGTT WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906		MALSGNCSRYYPREQGSAVPNSFPEVVELNV GGQVYFTRHSTLISIPHSLLWKMFSFKRDTAN DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRVPR ILVCGRISLAKEVFGETLNESRDPDRAPERYTS RFYLKFKHLMGAPASNFILGFWGLGQNQDK HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRWLGDPEHL
993	2343	A	8379		2794	MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIE DMVTTASTYLFEATEKRFFKNVSILIPENWK ENPQYKRPKHENHKHADVIVAPPTLPGRDEP YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDOPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRQRI VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVIEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQŁESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQITDLDATVHEDKIILTWTAPGD NFDVGKVQRYIIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPT PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
995	2345	A	8390	194	3421	PSSWDYRACLS AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				Sequence		DFLSMKQSPALAPEERCRRAGSPKPVLRADD NNMGNGCSQKLATANLLRFLLLVLIPCICALV LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYH ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENL FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG
	•		-			DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGR
						TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYITGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN
996	2346		8392	199	3085	KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSRSRSPSPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence		e .	914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE DDESGTENREEKDNIQPITE
997	2347	A	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS KQENLESLEATGREEEGGQGPPVTTKGVLLA LLMAGLALQPGTALLCYSCKAQVSNEDCLQ VENCTQLGEQCWTARIREWGDDSRQA
998	2348	A	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN MSDPRRPNKVLRYKPPPSECNPALDDPTPDY MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ NPQPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL VSKKANQLNFTEAKEACRLLGLSLAGKDQ VETALKASFETCSYGWVGDGFVVISRISPNPK CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE VFMETSTMSTETEPFVENKAAFKNEAAGFGG VPTALLVLALLFFGAAAGLGFCYVKRYVKAF PFTNKNQQKEMIETKVVKEEKANDSNPNEES
1000	2350	A	8406	2	777	KKTDKNPEESKSPSKTTMRCLEAEV KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT
·						AAIFTÄDGNMISASTLMDILLMNDFKLVINKI AYDVQCPKREKPSNEHTAEMEHMKSLVHRL FTILHLEESQKKREHHLLEKIDHLKEQLQPLE QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA LAWLTWWYSWDIMEPVTYFITFANSMVFF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ HFDVQQYNKLKEDLAKAKESLKQARHSLCL QMQVEELNEKN
1001	2351	A	8410	1400		VGFWERPLRSSRWFRRSLRRWEMLARAARG TGALLLRGSLLASGRAPRRASSGLPRNTVVLF VPQQEAWVVERMGRFHRILEPGLNILIPVLDR IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV LYLRIMDPYKASYGVEDPEYAVTQLAQTTM RSELGKLSLDKVFRERESLNASIVDAINQAAD CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE RRKRATVLESEGTRESAINVAEGKKQAQILAS EAEKAEQINQAAGEASAVLAKAKAKAEAIRI LAAALTQHNGDAAASLTVAEQYVSAFSKLA KDSNTILLPSNPGDVTSMVAQAMGVYGALT KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS
1002	2352	A	8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIHNKHAG
1003	2353	A	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

	000 10			The state of	I D 111 1	A Alexandra
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ĺ	in		location	
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ŀ	1	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	]	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
			<u>l</u>	sequence		nucleotide insertion
			}		}	SLDTENIDEILNNADVALVNFYADWCRFSQM
	ļ		į.	i '		LHPIFEEASDVIKEEFPNENQVVFARVDCDQH
		1	i		Ì	SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ
	1		Į.		1	RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS
	i	ļ	1			KRNIIGYFEQKDSDNYRVFERVANILHDDCAF
		1	ĺ	[		LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY
		l		1	ļ	LGAMTNFDVTYNWIQDKCVPLVREITFENGE
	ļ	1	1		j	ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL
		1	1	1		ISEKGTINFLHADCDKFRHPLLHIQKTPADCP
	1			1		VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL
		ļ		Į		HSGKLHREFHHGPDPTDTAPGEQAQDVASSP
						PESSFQKLAPSEYRYTLLRDRDEL
1004	2354.	A	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM
			1	1	1	ACAAARSPADQDRFICIYPAYLNNKKTIAEGR
	l	l .		l		RIPISKAVENPTATEIQDVCSAVGLNVFLEKN
	1	1		Ì		KMYSREWNRDVQYRGRVRVQLKQEDGSLC
	Į			ł		LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA
				1		DOSLOGGEGSKKGKGKKKKK
1005	2355	A	8453	90	530	OSHETKMOSGTHWRVLGLCLLSVGVWGQD
		1	"""	1		GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE
	}	.]	j .	1		ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF
	ŀ					SELEQSGYYVCYPRGSKPEDANFYLYLRARG
	1	1			1	NPGLONRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVORIRHEMNIFRLTGDLSHLAAIVILLLKIW
1000	2330	1 11	10,20	1	1 20,	KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS
		1		i		LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC
	j	1	1	1	ļ	OLCIFN
1007	2357	A	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL .
1007	2337	^	0433	1 43	333	SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT
	}	1	ŀ	}	ł	VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM
		1				GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA
	İ					FWWHNKGLALIFCILQSLALTWYSLSFIPFAR
1	1	ł	1	l	İ	DAVKKCFAVCLA
1000	0260	ļ.,	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP
1008	2358	A	8462	467	130	PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS
ł	ł	1			1	DPRWGCVGPSMPTSTCLPGAVEASTTKASLP
	1			1		KCPVDSSLPTPEACFL
1000		<u> </u>	-	124	054	
1009	2359	Α	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP
1	l	}	1		1	NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH
1		1	1	1		LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN
1		1	1	1	1	FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL
1				1	1	TKLLVHSSLVGSILSALSALVGFIILSVKQATL
					.]	NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD
1					1	CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL
ł		1	1			RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT
		<u> </u>			<u></u>	HDCGYEELLTS
1010	2360	A	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA
		1		1	1	HRVALCHLAGCQEQAAWYHTLQILFFLVSAY
	ļ.			1		FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT
						LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL
		1				SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT
		1	1	1		GTLETQFTCPFCNHEKSCDVKMDRARNTGVI
}	1	1				SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG
	1	1	1	1		PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC
		1				RAGFQCQ
1012	2362	A	8481	2810	1652	RTSTOKWOSVFNDSOEHLERFYCNPENDRM
1012	2302	1	0.701	2010	1002	RMKYGGQEFWADLNAMNVYETTEFDQLRR
	1	1	1	1	1	
Į			1	1	1	LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SVIRLIEEANSRGLKEVRFMMWNNHYILHNS FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR ILQILRVQNQFLWEKYKRKEYMNRKMFGR DRIINERHLFHGTSQDVVDGICKHNFDPRVCG KHATMFGQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI OYEEVSNTVSI
1013	2363	A	8488	2	517	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	Α .	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE
1015	2365	A	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCLLSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIATYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT
1016	2366	A	8511	1	453	KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN
1017	2367	A	8513	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

PCT/US01/03800

					N 11-1-1	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide		I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	to last amino	Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
	<b>i</b>		1	amino acid	of peptide	l=Inreonine, v=Vaime, w=Itypiopitali,
	1		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		1	peptide	i	/-possible nucleotide deletion, \-possible
	Ì		ļ	sequence		nucleotide insertion
						SLALGWEAFHALQILGFLILLIGTALYNGLHR
	İ	l	1	1		PLLGRLSRGRPLAEESEQERLLGGTRTPINDA
	1	] ·	1	1	· ·	S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV
1010	2500	^	1 0510	327	1	VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA
	1	1	1	1	1	YQLYQDPRNVWGFLAATSVTFVGVMGMRS
		1	İ	]	1	YYYGKFMPVGLIAGASLLMAAKVGVRMLM
	l .	1	1	ļ		TSD
		<u> </u>		ļ <u></u>	1000	VSAAAVNMEPPDAPAQARGAPRLLLLAVLL
1019	2369	Α	8526	2	1787	AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT
		1	1	1	1	MANIPUNGATION A PINET TO BEGAD DOLASA EMO
	1		1	1	1	PTGTHDHYMLEWFLTDRSGARPRLASAEMQ
		1	J	1		GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA
		1	1	1	1	QVGDERDYVCVVRAGAAGTAEAAARLNVF
		1		1	1	AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN
		1		1	1	GNPAPKITWYRNGQRLEVPVEMNPEGYMTS
	1	1	1		1	RTVREASGLLSLTSTLYLRLRKDDRDASFHC
		1			}	AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ
	i		1 .	į		FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP
		1		i	Ì	EYTLFRLODEOEEVLNVNLEGNLTLEGVTRG
		ļ	ļ	i		QSGTYGCRVEDYDAADDVQLSKTLELRVAY
	ļ	Í	i			LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP
	ļ	1	1	1	i	ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC
1	1	1	1			EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP
	1	1	ŀ			KADGSWREGDEVTLICSARGHPDPKLSWSQL
1		1	1	1		GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI
1	Į	1	1			SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV
	i i	1	1		1	SCEASNPHGNAKHVFHFGIVSIQISQAGVAV
1	1		1			MAVAVSVGLLLLVVAVFYCVRRKGGPCCRQ
į.		1			i	RREKGAP
1020	2370	A	8530	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS
1020	1 / (	,			ł	SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG
ļ	1	1	- (	ł		VLFSTGLLGNLLALGLLARSGLGWCSRRPLR
1	1	ļ	İ	1	1	PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA
	i i	1	1	ì		YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL
ł		-	- (	l .	1	SSTLQLLAMALECWLSLGHPFFYRRHITLRLG
	1	1	ļ	ì	1	ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG
		1	1	1	1	TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV
1	1	1		1		LATVLCNLGAMRNLYAMHRRLQRHPRSCTR
1	1	1	{	1		DCAEPRADGREASPQPLEELDHLLLLALMTV
1						LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA
	1	1	1	1		PUBLICATION ALL WALL TOWNED A WEST ALL SEEN
	1		1	[	1	EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI
			L			FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR
		-			1	KHDSGAADLERVTDYAEEKEIQSSNLETAMS
1		1		1		VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRRMEAVVFVFSLLDCCALIFLSV
1042	2312	^	3337	1	1	YFIITLSDLECDYINARSCCSKLNKWVIPELIG
		Į	1	1	1	HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI
		i	1	1		MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI
İ	1	1	ļ	1	1	KLGFHLLCFFMYLYSMILALIND
L						RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL
1023	2373	A	8540	26	431	MANANCE AND THE APPEAR A SOUTH A ACT OF THE SOUTH A
1		-	1	1	1	SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK
1		l	1	1		KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL
1		1	1		ì	NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS
		1	1			TAEALQL
	2374	HA	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT
1004	1 / 5 //1	l W	10344	[ *'3'	1	AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL
1024	23,4					
1024	23,4	1		İ	}	SALL SA AFLL VRKL PPLCHGL PTOREDGNPCE
1024	. 25,4		}			SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD
1024	. 23/4					SALLSAAFLLVRKLPPLCHGLPTQREDGNPCL FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT WIVEFFANWSNDCQSFAPIYADLSLKYNCTG
1025	2375	A	8546	2194	1707	LNFGKVDVGRYTDVSTRYKVSTSPLTKQLPT LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS TPTTVSDGENKKDK TVSFHKTMASLKCSTVVCVICLEKPKYRCPA
						CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS ALPTKTVKPVENKDDDDSIADFLNSDEEEDR VSLQNLKNLGESATLRSLLLNPHLRQLMVNL DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES
1026	2376	A	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS YAWANFTILALGVWAVAQRDSIDAISMFLGG LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL SLLLKPLSCCFVYHMYRERGGELLVHTGFLG SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY
1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK
1028	2378	A	8569	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLL LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFFYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY
1029	2379	A	8572	1	578	AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV
1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF
1031	2381	A	8580	905	340	RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
1032	2382	A	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

			Lorio		Descripted and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	]	in	nucleotide		I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	ļ	USSN	location	corresponding	M=Methionine, N=Asparagine, P=Proline,
-pas	uence	1	09/496	correspondi	to last amino	
uence		Į.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ	Í	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ı	j	}		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1		peptide		/=possible nucleotide deletion, \=possible
	ľ			sequence		nucleotide insertion
						QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD
	İ	1	1		l	VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
		ì	1		l	TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV
		1	1			MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
		Į				GERAFMNRTVGEIMWGYKDPLVNLINKYFP
	1	1	j	ì	ļ	GMFPFKDKFGLFAELNNSDSGLFTGFTGVQNI
		1		ļ	1	SRIHLVDKWNGLSKVDFWHSDQCNMINGTS
ŀ				1	· ·	GOMWPPFMTPESSLEFYSPEACRSMKLMYKE
l	1		j	1	1	SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
				1	ļ	CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL
i	1	1	1	ì	ļ.	AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV
]			j	ì	ł	KLOLSLYMKSVAGIGQTGKIEPVVLPLLWFA
	1		1			ESGAMEGETLHTFYTQLVLMPKVMHYAQYV
	1	1	1			LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK
	ì	<b>\</b>			i	GSKDKEAIOAYSESLMTSAPKGSVLQEAKL
				+	-	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
1033	2383	Α	8595	595	767	FCLLLSLVSSSLVSLSLCPPLTQA
L				1	1	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ
1034	2384	A	8597	640	164	PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL
į.	1	1	.			PMMQ1GQK1CMDPAVIAGVLSRRSPGDAIL
l .	1	1 .	1			VNMGDRTSMVQDPGSQAPTSWISESQVFQTT
1	į.		ł	i		EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG
	1	1		}	j	GAGYVRSSQDLSCDFCNDVLARAKYLKRHG
	1	ĺ	1		l	F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL
1					1	SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG
1		ſ	1		1	SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI
1	ŀ	1	1		1	HVYKKNGVGKVGDQILLAIKGQKKKALIVG
1	1		1		1	HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI
	ł	1	1			KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	+1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV
1030	2360	1	1 0000	1 -		TQYLQPRSPEECKMFACAKLACTPSLIRAGSR
1.	1			1	]	VAYRPISASVLSRPEASRTGEGSTVFNGAQNG
i		1				VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG
1	1		İ	İ	1	VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY
		1		ļ		AILGFALSEAMGLFCLMVAFLILFAM
1000	- 0200		9616	12	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT
1037	2387	A	8615	12	2307	GMVAHINNSRLKAKGVGOHDNAQNFGNQSF
ł	1	1	1	1	1	EELRAACLRKGELFEDPLFPAEPSSLGFKDLG
ì						PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI
ĺ	I	1				CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG
ł		ļ	l	}	ļ	QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL
<b>\</b>		1	I			PTKNDKLVFVHSTERSEFWSALLEKAYAKLS
1	1	1		1		
1		1				GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP
1				}	1	PONLLRLRKAVERSSLMGCSIEVTSDSELES
1					İ	MTDKMLVRGHAYSVTGLQDVHYRGKMETLI
				1		RVRNPWGRIEWNGAWSDSAREWEEVASDIQ
1				1	1	MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL
1		İ				TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC
		- 1	1	1		RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV
1		1		1		VVCTCLVALMQKNWRHARQQGAQLQTIGFV
[		]	}	J .	}	LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI
1			- 1	1		FTNSREVSSQLRLPPGEYIIPSTFEPHRDADFL
		1				LRVFTEKHSESWELDEVNYAEQLQEEKVSED
J			1	1	1	DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR
1		1				MAIKFKSFKTKGFGLDACRCMINLMDKDGSG
		l l				KLGLLEFKILWKKLKKWMDIFRECDQDHSGT
		- 1		1	l	LNSYEMRLVIEKAGIKLNNKVMQVLVARYA
1		- 1		,	<b> </b>	DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT
	- 1	ļ	1			GHICLSLEQVLGEGWEGICRIAPACPSTPPPPS

SEQ ID	SEO ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
uence		1	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
donoc		i	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	•			peptide	sequence	/=possible nucleotide deletion, \=possible
Į		ļ ·		sequence		nucleotide insertion
		<del> </del>	<del> </del>	Sequence		SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI
						EAL
1038	2388	<del>                                     </del>	8621	3	1494	
1038	2388	A	8021	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY
1					i	HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL
			1		}	ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF
1		İ		ĺ		GGIETLRVPSELVWLPEIVLENNIDGQFGVAY
		[	1	ł		DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF
1	i	1		İ	1	PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG
,	J	j	]	ļ	ļ	KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH
					Ì	GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV
ļ	Ì		Ì	}	[	LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT
1	`.	1	1	l		VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI
<b>J</b>	1	ļ	1	]	}	VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL
1		1			[	LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE
1		]				ELILKKPRSELVFEGQRHRQGTWTAAFCQSL
1	1	i		1	ļ	GAAAPEVRCCVDAVNFVAESTRDQEATGEE
İ		1				VSDWVRMGNALDNICFWAALVLFSVGSSLIF
					<u> </u>	LGAYFNRVPDLPYAPCIQP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL
:	}	l				QAMENELPVPHTSSSACATSSTSGASSSSGCN
ļ				1		NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ
1	1	1		ļ	1	ALHRQPSTAAQYLQQMYAAQQQHLMLQTA
1	ļ			ļ	1	ALQQQHLSSAQLQSLAAVQQASLVSNRQGST
J	1			}	}	SGSNVSAQAPAQSSSINLAASPAAAQLLNRA
1						QSVNSAAASGIAQQAVLLGNTSSPALTASQA
l	Į		ļ	l .		QMYLRAQMLIFTPTATVATVQPELGTGSPAR
1	l	}		1	·	PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL
1					1	PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF
		1				HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPO
1	l	l		ł	ļ	RRDSFSGVKDSNNNSDGKAVAKVKCEARSA
	[			1	ļ	LTKPKNNHNCKKVSNEEKPKVAIGEECRADE
1		į.			ļ	QAFLVALYKYMKERKTPIERIPYLGFKQINLW
	ł	İ	i	1		TMFQAAQKLGGYETITARROWKHIYDELGG
1	l	!			•	NPGSTSAATCTRRHYERLILPYERFIKGEEDKP
	İ					LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP
1				1		KSKKEKENAPKPQDAAEVSSEQEKEQETLISO
1	l		i	l		KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD
				1		PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA
1	1					
		Į.	1	1		PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWOTR
	1			ļ		MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSOTIYP
1441	20,51	1	0040	113	1472	
	1	1	[	1	1	GIKARITQRALDYGVQAGMKMIEQMLKEKK
Ì				1		LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP
}	1	1	j	J		NTSLAFVPGVGIKALTNHGTANISTDWGFESP
	1			ł	]	LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK
				1	Ì	ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE
		1	1	ļ.		NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER
	}		j	j		SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS
1				1		TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM
				ļ.		VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK
1	1	l		ł		NSTVETIVSMDFVASTSVGLVILGQRLVCSLS
		1	1	ŀ		LNRFRLALPESNRSNIEVLRFENILSSILHFGVL
	1			Į.	1	PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF
1			1	I		LLISTDLKYETSSKQQPSFHVWEGLNLISRQW
		<u> </u>			ļ	RGKSAP
1042	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR
		_		ļ.,		LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM
	<del></del>				<del></del> -	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	,,,,,,,	in in	nucleotide	location	F=Phenylelanine, G=Glycine, H=Histidine,
		i '	USSN	location		
eotide	seq-				corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
<b>i</b>		İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
t				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ		1	į	peptide		/=possible nucleotide deletion, \=possible
L			<u> </u>	sequence		nucleotide insertion
						TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT
						LDMITSTDHVLEQDFWICFTFYSVKERQI
1043	2393	Α	8688	359	17	GLKTRAPATPTFQREVLGPAKQDMQRRCPRI
	į			ĺ		GLMTSLLKPIKRRWRDYKRWKSGGFTGESC
1	İ	ļ	l			HHADTLGDRGGLOGDHSELLOWOKRILRTE
			1			GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS
						GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS
			ł			YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL
			1			VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL
				1	1	NLALADLLFALTLPIWAASKVNGWIFGTFLC
	1		Į.			KVVSLLKEVNFYSGILLLACISVDRYLAIVHA
1	1	1	١ .			TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR
l.		Į.				RTVYSSNVSPACYEDMGNNTANWRMLLRIL
1	ļ		ľ	1	1	
			İ			POSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK
	}		ļ	1		HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM
		i	ł		l	RTQVIQETCERRNHIDRALDATEILGILHSCLN
					ļ	PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS
-		<u> </u>	-	L		RPSFVGSSSGHTSTTL
1045	2395	A	8724	254	3184	FRANLAITVANRRGAQGGKMHTCCPPVTLEQ
	ł		]	ł		DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY
i .	1	l	ŀ			YGEICDNACPCEEKDGILTVSCENRGIISLSEIS
		İ				PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL
		1	ļ			HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL
j	}	ļ	J .	ļ	ļ	ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF
						GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL
Ì		İ				DLRGNRLKLLPYVGLLQHMDKVVELQLEEN
			1	ļ		PWNCSCELISLKDWLDSISYSALVGDVVCETP
i	ĺ		Į.	İ	ĺ	FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP
1	ļ		İ			LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP
						KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA
	i					YQTKSPVPLECPTACSCNLQISDLGLNVNCQE
1	ļ		1			RKIESIAELQPKPYNPKKMYLTENYIAVVRRT
i	ĺ	l	1	ŀ		
						1
1	•	1		ļ		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS
					·	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGGSSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP
				_		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQQPPPQLQLQPGEEERRESHHLRSPAYS
				_		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH
				-		DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE
1046	2306		9732	-	452	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEEERRSHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW
1046	2396	A	8736	28	452	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC
1046	2396	A	8736	28	452	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLAGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEUSGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA
1046	2396	A	8736 8741	28	452	DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA ALPGTPQQTVILNTDGKVKSFTSPHSNPNLPP
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLAGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEUSGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA ALPGTPQQTVILNTDGKVKSFTSPHSNPNLPP
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQOPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		VAVPNGOPPSAARYMPREVPPRFRCQQDHK  VALKRGQPPPPSCMLLGGGAGPPPCTAPGAN PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWGSGASSNNGTSPNPHHWDKVIVDGS DMEEWPCIASKDTESSSENTTDNNSASNPGSE KSTLPGSTTSNKGKGSQCQSASSGNECNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGPGSGFSNPNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGQTSREQQSKME NAGVNFVVSGREQAQHINTDGPKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNRSTGGSWNFGPQDSNDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWO EGGQNEIGTWGNGGNASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQPPQPPPPQ PEASGSWGGPPPPPGNVRPSNSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFFNSGLSP GLFNVGPQLSPQQIAMLSQLPPIQFPGLACQL LLQQQQQQLLQNQRQRGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESFRQWTSMME GLPSVATQESPQUIAMLSQLPPIPGTPNK IGSKSSNASWPPEFQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNFKFSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSTFLAQAPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNSSSL
1049	2399	A	8748	200	1387	PHRMGSPAPLLPGDLLGGGSDSI  VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA PFALSALLYGANNNLVTYLQRYMDPSTYQVL

SQ ID NO: of mucle colide sq- usince of the control	<u> </u>	l one m	14.4	CEC	Dung! etc. 1	Daniel Service 1	LANDS AND COMMENTS OF THE COMENTS OF THE COMMENTS OF THE COMMENTS OF THE COMMENTS OF THE COMME
nuclectide seq uence USSN 09496 location corresponding profession of the control	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence unce sequence s		1	noa				
sequence    914   914   19   19   19   19   19   1			İ				
1914   ng m first simino acid residue of peptide sequence   peptide				I .			
mino acid residue of peptide requence periode sequence peptide sequence peptide sequence periode sequence periode sequence periode sequence periode sequence periode sequence periode sequence s		uence	1				
Presidue of peptide sequence	uence			914			
peptide sequence    Possible nucleotide deletion, \( \sigma \) possible nucleotide insertion   SNLKIGSTAVLYCLCERRILS/VROGLALLIL     MAGACYAAGGLQ/VGOTITESPIPAAASS     MPHHTPI-GILLILILYCLISGISS-VYTELIAMKR     QRLPLALQNLFL-YTEGYULIN-GILHAGGSGP     GILBGFSGWAALVVLSQALNGILMSAVMKH     GSRTILLIGLAMKL-YYGSR     GSRTILLIGLAMKL-YYGSR     WVSSMGFEELLEQVGSCLY-VAVLSAVLLALD-ALALP     RYLLPLIFILLIPILAAVPARRCALPA-NATS-ALALP     RYLLPLIFILLIPILAAVPARRCALPA-NATS-ALALP     NTTI-GERCSRGELEDEPAT-YPCSQQWETOH     THE SERSITATESQWDLIV-CORGIGNIFAASTIFFA     GVLVQAVAFOTL-SDEFGRRRILLIV-AYVSTLV     LELASAVSVYMFAITTRIT-GSSALAGFTIV     MPLELEWI-DVERRITV-AGVLSSTFWTGGYML     LALVQVILRDWR WLLLAVTI-CARGISLS W     VVSSARWILLT-QGHVER-HRYLL-CARLINGR     PVCEDS-SSQ-AVSKV-AAGER-VYRR-SYLDLF     RTPRLRHISL-CVV-WFGVNFSYYGLSLD-VS     GGLNV-VYTGDL-FRAVEL-STATL-YSST-SYLD-STATL-SYLD-SYLD-STATL-SYLD-SYLD-STATL-SYLD-SYLD-SYLD-SYLD-SYLD-SYLD-STATL-SYLD-SYLD-SYLD-SYLD-SYLD-SYLD-SYLD-SYL			1	ĺ			
				Į.		sequence	
SNLKIGSTAVLYCLCERRISYROGIALILLI MAGACYAGGLQVPONTIPSPPPAAAAS MPHHTPI GILLILILYCLISGI SSVYTELIMKR QRLPLAQUELT YTFOVLINLIGHAGGSOP GILBGFSGWAALVVLSQALNGILMSAVAKH GSSTELFVYSCSLVVAVLSAALLRQLTAA FFLATLLIGI AMRLYYGSR HQDVWLSAHLFREYDGTLSSCLRFAYPQALD SPEATH THE STANDAV ALL ALP FRULPHPLE PIFLAAVPAHRCALPGAPANES HQDVWLSAHLFREPDGTLSSCLRFAYPQALD SFISSTIATESQWDLVCEQKGINRAASTFFFA GULVGVAVFGYLSDFGRRELLLAYAVSTLV LGLASAASVSYMFAHTRILTGSALAGFTHV MPLEE JERUDVERTVAGVISSTFTYTGGVML LALVCYULRDWRWLLLAVTLFCAFGILSLWW VESBAWRLTQGHYSGAVSKVAAGGRVVRRPSYLDLF STPRLKHRISLCCVVVFGVNFSYNGGSLDVS GIGINVYQTOLLFGAVELHSKLLVYLSVRSYA GRRITQAGTLLGTALAGFTRLUVSSDMSWS TVLAVMGKAFSEAAFTTAVLFTSELYFTVLR GRRITQAGTLLAGTALAGFTRLUVSSDMSWS TVLAVMGKAFSEAAFTTAVLFTSELYFTVLR QTGMGLTALVGRLGGSLAPLAALLGGWLS LFRUTYGGIALLAGGTLGUSJAHCAALLGGWLS LFRUTYGGIALLAGGTLGUSPAGVENDE EDGGVVV VEKELBERPMKQVQN 1051 2401 A 8759 515 1625 ERTPVAVSSAPSGDSEGDEETTGDEVSSHTS ERTPVAVSSASFSGDSEGDEETTGDEVSSHTS TVLAVMGKAFSEAAFTTAVLFTSELYFTVLR FOLDENSTRIST FRUTYGGIALLAGGTLGUSPAGNAV VOHTAAVVSAKSYMCVGGRALSSPSGLGR HLIHSEDQRSNCAVCGARFTSHATTNSEKLP EDGGVVV VEKELBERTREPVGGVVEHEV TGNINSPPLLELCCCPLCQLDCGSREQILAHV VOHTAAVVSAKSYMCVCGRALSSPSGLGR HLIHSEDQRSNCAVCGARFTSHATTNSEKLP EVLNMESLFTVHNIGGPSSAGKGLAFSFONSSH ONEPLEVPLQRLIEBERTAKKSRRNDFTEPERE VRMMDREAKRLORMGETDEGRARRLIGRR ONEPLEVPLQRLIEBERTAKKSRRNDFTEPERE VRMMDREAKRLORMGETDEGRARRLIGRR PCEPURQUERERTAKKSRRNDFTEPERE VRMMDREAKRLORMGETDEGRARRLIGRAVLIGUV ATVRSLGLIVPFYTHRAGFTVAVLAGMALGIGV VPSEMCHALLGLIVL ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUV ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUV ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUV ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUS LICEBOTYVVALAVVALAGLIGUV ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUS LICEBOTYVVALAVVALAGLIGUV ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUN ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUN PDBWGFPRONTAVILLIGUN ATVRSDLSTFIILLAYSOYEVVOMILSVLIGUN VINDEAHLVKANINGGCKOKNINVINVYTIFV VALKACIQUMKKREEREMKKKRRNDFLISV SALKKVENMSSNQOLONDSDEFM VINDEAHLVKANINGGCKEREMKKKRRNDFLISK VANELLBRIVVENELSALAKEEREMKKKRRNDELISV SALKKVENMSSNQOLONDSDEFM  1054 2404 A 8769 344 527 8FALTLLICALACRISCOVPSECSGLG		ļ		1			
MAGGACYAAGGLQVPGNITISPPPAAAASP MPHHPIGLILLII, YCLISGISSYYTELIMKR QRLPLALQNLFLYTFGVLINLGHAGGGSQP GILEGFSGWAAL VYLSQALNGLIMSAYMKH GSSITRLEVYSCSLVVNAVLSAMVLRLQLTAA FPLATLIGLAMRL YYGSR  1050 2400 A 8758 3 1669 WYSSNGFEELLGQVGGFGPFGLRNYALLALP FRATLIGLAMRL YAPARCALPOAPANES HQDVWLEAHLPREPDGTLSSCLRPAYPQALP NTILGERQSRGELEDEPATYPCSQQWEYDH SEFSSTIATESQWDLVCEQKGINRAASTFFA GVLVQAVAFGYLSDRFGRRELLLAYYSTLV LGLASAASVSYVMRATRITLGSALAGFIIV MPLELEWLDWERRTVAGVLSSFFWTGGVML LALVCYLIRDWRWLLAVTLPCARGLISLWW VYESARWLITQHIVECHARTLUSSDMKSWS TYLAVMGKAFSEAAVGERVVRRPSYLDLF RYPELRIISLCCVVVWFGVNFSYYGLSLDVS GLGINVYQTQLLFGAVELPSKLLYYJSWYA GRRITQAGTILGTALAFGTRILVSSDMKSWS TYLAVMGKAFSEAATTTALFTSELYFTVLR QTGMGLTALVGRLGGSLAPLAALLDGWWLS FYRTYGGIALAAGGTALLPETRGAQFET QDVEKSAPTSLQEEBMPMKQVQN  1051 2401 A 8759 515 1625 ERTPYAVSSAFGOSEGDEETTODEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV YOHTAAVVSAKSYMCZVCGRALSSPGLIAHV YQHTAAVVSAKSAMCZVCGRALSSPGLIAHV YQHTAAVVSAKSAMCZVCGRALSSPGLIAHV YQHTAAVVSAKSAMCZVCGRALSSPGLIAHV YQHTAAVSASSPLIAHQUSTATACKSCAMVALIGA GREEGEAACKARRAAALAGANSTPT AGRICGRATHACKACAACAACAACAACAACAACAACAACAACAACAACAA					sequence		
MPI-HITPLGILLLILYCLISGLSSVYTELLMKQ   QRLPLALQNIF_YTGVILNAGLAGGGGG   GLLEGFSGWAALVVLSQALNGLLMSAYMKG   GSITRI-FVVSCSLVYMAVLSAVLRI_QUTANA   GSITRI-FVVSCSLVYMAVLSAVLRI_QUTANA   FFILATILICILAMRLYYGGR   PROSERVE	,						
1050   2400   A   8758   3   1660   WVSSMGFEELEQVGGGFGFQLLSAAVALLALQLTAA   FILATILICIAAMELYYGSR   WVSSMGFEELEQVGGGFGCLRNYALLALP   RVILIPLIFILAAVAHRCALPGAPANFS   HQDVWLEAHLREPEDGTISSCLEPAYPQALF   NTTLGEERGSGGELDEPATVPCSQGWEPTOL   SEPSSTIATESQWD LVCERGGINRAASTFFFA   GYLVQAVAFGYLSDRFGGRELLLVAYVSTIL   LGLASAASVSYVMFAITRILTGSALAGFTILV   MPLELEWLDVEHRTVAGVLSSTFWTGGWM   LALVQYILRDWRYMFAITRILTGSALAGFTILV   MPLELEWLDVEHRTVAGVLSSTFWTGGWM   VFSSARWLTQGIVKZAHRYLLHCARINGR   PVCEDSTSQBAVSKVAAGERVLYJSWYA   GREILTGAATELTSTELLYFTVSL   GIGLINVYQTQLEGAMELPSTLLVSSYMSKSW   TVLAVMGKAFSEAAFTTAYLTFSLLYFTVSL   RTFRLRHISLCCVVVWFGVNFSYYGLSLOW   VFSSARWLTQGIVKZAHRYLLHCARINGR   PVCEDSTSQBAVSKVAAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSW   TVLAVMGKAFSEAAFTTAYLTFSLLYFTVSL   PVCEDSTSQBAVSKVAAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSW   TVLAVMGKAFSEAAFTTAYLTFSLLYFTVSL   PVCEDSTSQBAVSKVAAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSW   TVLAVMGKAFSEAAFTTAYLTFSLLYFTVSL   PVCHORGATALVGRIGGSLAPLAALLDGVUS   PVCEDSTSQBAVSKVAAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSW   TVLAVMGKAFSEAAFTTAYLTSELLYFTVSW   PVCEDSTSQBAVSKVAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSWAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSWAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKSWAGERVLYJSWYA   GREILTQAGLFTGAVELPSTLYSSYMSKATA   PVCENTRATE   PVC		1	1	Į.		}	
		ł	İ	ì			MPLHITPLGLLLLILYCLISGLSSVYTELLMKR
							QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP
		ŀ	1	Į.			GLLEGFSGWAALVVLSQALNGLLMSAVMKH
1050   2400   A	1		j	1	ŀ	)	GSSITRLFVVSCSLVVNAVLSAVLLRLQLTAA
RVLLPHELLPIFLÄAVPAIRICALPGAPANFS HQDVWLSHIPRSPOSTISSERAYPOALP NTTLGEERQSRGELEDPATVPCSQGWEYDH SEPSSTIATESQWDLVCEQKGINRAASTFFA GVLVGAVAFGYLSDRFGRRELLIAVATVSTLV LGLASAASVSYVMFAITRILTGSALAGFIIV MPLELEWLDVEHRITVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLFCAFGILSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW VPESARWLLTQGHYKEANFYLHCARIGLSLWW GRAITQAGTLLGTALAFGTRILLVSLSWYA GRRITQAGTLLGTALAFGTRILLVSLSWYA GRRITQAGTLLGTALAGTALLIPFIRQAQLPETI QTGMGLITALVGRLGGSLAPLAALLIDGWUS LPELTYGGGALLAAGTALLIPFIRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN TQTGMGLITALVGRLGGSLAPLAALLIDGWUS LPELTYGGGALLAAGTALLIPFIRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN TGMGLITALVGRLGGSLAPLAALLIDGWUS LPELTYGGGALAAGTALLIPFIRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN TGMGLAAGTALLIPFIRQAQLPETI QDVERKSAPTSLQEEEMPMKQVGN TGMLSAPGGSLGSEQLIAHW YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSNCAVCGARFTSHATTNSEKLP EVLNMESLFTVHNEGPSSAGEGDLAFSPFVYP AGBLLVCNNCAAYRKLLEAQTFSVRKWALRR QNEPPLEVRLQRLEREATAKKRRADNETFEERG VRRMRDREAKRLORMFTEPKQARLINGTSFVRKWALRR QNEPPLEVRLQRLEREATAKKRRADNETFEERG VRRMRDREAKRLORMFTEPKQARLINGTSFVRKWALRR GNEPPLEVRLQRSLLGKMAFTERKAKLKRR LEKMDMMLRAQFGGDPSAMAALAAEMNFF QLPVSVVELDSOLLIGKMAFTERKAKLKRR LEKMDMMLRAQFGGDPSAMAALAAEMNFF QLPVSVVELDSOLLIGKMAFTERKAKLKRR LEKMDMMLRAQFGGDPSAMAALAAEMNFF QLPVSVVELDSOLLIGKMAFTERKAKLKRR LEKMDMMLRAQFGGDLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLEDDTSGYSSQFGGY PATGADVAFSVIHLLGDPMANVAMAYGSI ASKGKMYHRELIRFVSVSKUKSYFFAVDTA YVAKKLGLLVPPYTHQNWEVQYSRDAAPPP RQDLNAPDLYIPTMAFTIVYLLAGMALGIQK RFSFEVLGLCASTALVWVWEVLALLLGLYL ATVRSDLSTFFILLAYSGYKYKENDLIKHGW PEDIWFHVDKLSSAHVYLRIHKGENIEDIPKE VLMCAHLVKKYENPEDLIKHGW PEDIWFHVDKLSSAHVYLRIHKGENIEDIPKE VLMCAHLVKKYENPEDLIKHGW PEDIWFHVDKLSSAHVYLRIHKGENIEDIPKE VLMCAHLVKANIGGCKMYSNESLGACKPTVCSMP PEDIWFHVDKLSSAHVYLRIHKGENIEDIPKE VLMCAHLVKANIGGCKMNVTVYTVEK VNEUNILEKTKVERPFDLAABECECRDEER NEKKAQIQEMKKREGEBRKKKREMDELRSY SSLMKVENNSSNQDGNDSDEM			1				FFLATLLIGLAMRLYYGSR
HQDVWLEAHLPREPDGTISSCLREATPQALP NTTIGERGSRGELEDPATVPCSQGWEYDH SEPSSTIATESQWDLVCEQKGLINRA-STIFFA GVLVGAVAFGYLSDRFGRRRLLLVATVSTLV LGLASAASVSTVMRAITRILTGSALAGFIILV MPLELEWLDVEHRITVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLPCAFGILSLWW VPESARWLLTQGHVKEAHRYLLHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLE RTPRIRHISLCCVVWWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTGLAAGTALLPTSLCVFTYLOX GRICTORY G	1050	2400	Α	8758	3	1660	WVSSMGFEELLEQVGGFGPFQLRNVALLALP
HQDVWLEAHLPREPDGTISSCLREATPQALP NTTIGERGSRGELEDPATVPCSQGWEYDH SEPSSTIATESQWDLVCEQKGLINRA-STIFFA GVLVGAVAFGYLSDRFGRRRLLLVATVSTLV LGLASAASVSTVMRAITRILTGSALAGFIILV MPLELEWLDVEHRITVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLPCAFGILSLWW VPESARWLLTQGHVKEAHRYLLHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLE RTPRIRHISLCCVVWWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTQLIFGAVELPSKUVRRPSYLDLE RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINYYGTGLAAGTALLPTSLCVFTYLOX GRICTORY G			1	"	ł		, , , , , , , , , , , , , , , , , , , ,
NTILGERGSRGELEDEPTYPCSGGWEYDH SEPSSTIATESGWOLVCEOKGINRAASTFFFA GVLVGAVAFGYLSDRFGRRRLLLVAYYSTLV LGLASAASVSYVMFAITRILTGSALAGFIIIV MPLELEWLDVEHRITVAGVLSSTFWTGGVML LALVGYLRDWRWLLLAYTLPCAFGILSLWW VPESARWLLTQGHYKEAHRYLHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRILHISLCCVVWFGOTHALAGTILLTSALWYVFYSGLSLDVS GLGLNYVGTOLLFGAVELPSKLLVYLSVRYA GRRLTQAGTILGTALAFGTILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLETSELYPTYLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTALLTGFTRQAQLPETI ODVERKSATISLQEBETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSPLLELCCCPLCQLOGISEQLIAHV YQHTAAVVSAKSYMCEVCGRALSSPGSLGR HLLISEDORSNCAVCGARTAFTNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGBLLVCNNCAAYRKLLEAQTFSVRKWALRR QNEPELEVRLQRLEBERTAKKARTHNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGBLLVCNNCAAYRKLLEAQTFSVRKWALRR QNEPELEVRLQRLEBERTAKKARRDNETIEEBEE VRRMRDREAKRLQRMGETDEGRARALGRDR EAMMLKRANFTFERRQARLGRDR EAMMLKRANFTFERRQARLGRDR EAMMLKRANFTFERRQARLGRDR EAMMLKRANFTFERRQARLGRDR EAMMLKRANFTFERRQARLGRDR ASHVFYPTRRSGPSGTASVGNSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVAFPGGLGRYPGRGAA ASLVFYPTRRSGPSGTASVGNSSLH AGRICHMUNKELHEFVSVSKKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDALPP RQDLNAPDLYJPTMAFITYVLLAGMALGIGK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFIELLAYSGVKKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDALPP RQDLNAPDLYJPTMAFITYVLLAGMALGIGK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFIELLAYSGVKKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDALPP RQDLNAPDLYJPTMAFITYVLLAGMALGIGK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFIELLAYSGVKKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDALPP RQDLNAPDLYJPTMAFITYVLLAGMALGIGK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFIELLAYSGVKKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDALPP RQDLNAPDLYJPTMAFITYVLLAGMALGIGK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFIELLAYSGVKKYFFAVDTA YFTSSSVNSSAYTTYMGKKKYENEDLIKHGW YFTSSSVNSSAYTTYMGKKKYENEDLIKHGW YFTSSSVNSSAYTTYMGKKKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRIHKGENIEDIPKE VLMCAHLVKVRFPDLAARECTREERE NEKKAQIQEMKKREGEMKKKREMDELRSY SSLMKVENNSSNQDGNDSDEM			l		ļ		
SERSSTIATESQWDLVCQKGLINRAASTFFFA GVLVGAVAGVI.SDRGKILLIVAVYSTLV LGLASAASVSYVMFAITRTLTGSALAGFTIIV MPLBLEWLDVEHRTVAGVLSITFVGGVML LALVGYLIRDWRWLLIAVTLPCAPGILSI.WW VPESARWI.LTQGHVKEAHRYLLICARLNGR PVCEDSFSQEAVSKVAAGREVVRRPSYLDLF RTPRLRHISLCCVVVWFGYNFSYYJGLS.DVS GLGLNVYQTQLLFGAVEPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALDGUWLS LPKLTYGGIALLAAGTRALDETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVON  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS DVDVERKSAPTSLQEEEMPMKQVON  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVVKEKLENTEQAGLPETI VQHTAAVVSAKSYMCPVCGRALSSPGSLGA HLIHIBEDGRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAGGKDLAFSPVYP AGILLVCNNCAVTRKLLEAQTPSVKKWALRR QNEPLEVRLORLERERTAKKSRRDETFEERE VRRMEDREAKRLGNNGCTDEGRARLGRDR EAMRLKRANETPEKRQARLTREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QNEPLEVRLORLERERTAKKSRRDETFEERE VRRMEDREAKRLGNNGCTDEGRARLGRDR EAMRLKRANETPEKRQARLTREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QNEPLEVRLORLERERTAKKSRRDETFEERE VRRMEDREAKRLGNNGCTDEGRARLGRDR EAMRLKRANETPEKRQARLTREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QNEPLEVRLORLERERTAKSKSRNOSTETERER VRRMEDREAKRLGNNGCTDEGRARLGRDR EAMRLKRANETPEKRQARLTREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QNEPLEVRLORLERTEAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF PRODUNAPPPILTDITSGGYPGGGAA ASLVFVPTIRRSGPSGTASVAAMAYHSGYGA HGSKIRARAAPDPPPLFDDTSGGYPGGGAA ASLVFVPTIRRSGPSGTASVAAMAYHSGYGA HGSKIRARAAPDPPPLFDDTSGGYSQFGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHREVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPTHAHTYVLLAGMALGIGK RFSEVLGLCASTALVWVVMEVLALLIGLGLL LFGSDGYYVALAWTSSALMYTIVLALGHGLICH LFGSDGYYVALAWTSSALMYTIVLALGHGLICH LFGSDGYYVALAWTSSALMYTIVLALGHGLICH VTTFILVR SNLKKTADMDVGQIGFRRQKDVKTVTVEK VPHILNRLEKTKVVERPPJLAABEKCTNTSREEN NEKKAQIQEMKKREKEEMKKKREMDELRSY SNLKKYENMSSNODONDSDEFM NEKKAQIQEMKKREKEEMKKKREMDELRSY SNLMKYENNSSNODONDSDEFM NEKTARDDONGGRER NEKTARDDONGGGFRRNSCWFTSRESCLGACKETVCSMP	!	ł	1	i	İ		
GYLVGAVAFÖXLSPFGRRALLIVAYVSTLV LGLASAASYSYMMATTATIGSAL AGFTIV MPLELEWLDVEHRIVAGVLSSTFWTGGVML LALVGYLIRDWRWILLIAVTLFCAFGLISLWW VPESARWILTGGHIVAGGENURRYLLICARLNGG PVCEDSFSGEAVSKVAGERVYRRSYYDLS RTPFLARBISCCVVVWFGYNFSYYGLSLDVS GLGLNVYGTQLLFGAVELPSKLLVVJ.SVRYA GRRITQAGTLLGTALAFGTRILVSSDMKSWS TVLAVMGRAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRIALGAGSLAPLAALLDGWULS LPKLTYGGIALLAGGTALLYSDMKSWS TVLAVMGRAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRIALGAGSLAPLAALLDGWULS LPKLTYGGIALLAAGTALLPSTRQACPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGOBSEGETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSRGLIAHV VQHTAAVVSAKSYMCPVGGRALSSPGSLGR HLLHSEDQRSNCAVGGARFTSHATFNISELP EVIANGESLFTVHNEGESRGCHAFSPPVYP AGILLVCNNCAAVRKLLEAQTFSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPERERE VRRMBDRBAKRLQRNQETIDEGRARRLQRDR EAMRLKRANETPEKRQARLIFRERAKRLKRR LEKMDMMIRAQFGQDPSMAALAAENNFF QLPVSGVELDSQLLIGKMAFEEQNSSLH  1052 2402 A 8763 1106 70 REGHGGRORGGGGRVAPGGGGRYPRGGAA ASLVFVPTRRSGPSGTAŠVAAMAYHSGYGA HGSKHRARAAFDPPLFDDTSGGYSQTGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHRPVSVSKLKYFFAVDTA YVAKKLGLIVPYTHONWEVQYSRDAPLPP RQDLNAPDLYTPTMAFITYVLLAGMALGIGK RFSFEVLGLCASTALVWVVMEVJALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMLSVLTGL LFGSBGGYVVALAWTSSALMYFITYSLRTAAL GPDSMGGPVRQRLQLYLTLGAAAFQPLITY WLTFHLVR  1053 2403 A 8768 2 712 RPPRWYPELRELSAAAPRWSHRTAPGIMVF YFTSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHIVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSINGGCKMNNVNVYTEW SNLKKTADMDVGQIGFHRQKDVXTYVEK VMELNRLEKTKVERPPDLAAEKSCRDRER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM NEKKAQIGEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM NEKKAQIGEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM	Ì	1	ļ	İ	Ì	,	
LGLASAASVSYVMEATRITLTGSALAGFTIIV MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWILLAVTLPCAPGILSLWW VPESARWILITQGHVKEAHRYLLICARINGR PVCEDSFSQEAVSKVAAGERVYRRFSYLDLF RTPRLRHISLCCVVVWFGVNFSYYJCLSLDVS GLGLNYVGTQLTGAVSEPSKLLVYLSVRYA GRRITQAGTILIGTALAFGTRLLVSSMKSWS TVLAVMGKAFEAAPTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAGTALLDETRQAQLPETI OTSMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAGTALLDETRQAQLPETI OTSMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAGTALLDETRQAQLPETI OTSMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAGTALLDETRQAQLPETI OTSMGLTALVGRLGGSSADEGETTQDEVSSHTS EEDGGVVVKEKELENTEQGTSEGLIAHV VQHTAAVVSAKSYMCPVGGRALSSPGSLGR HLLHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMGSLFTVHNEGPSSAGGGLIAFSPFVYP AGILLVCNNCAAYRKLLEAQTFSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDETFERER VRRMEDREAKRLQRNGTDEGRARRLQRDR EAMRLKRANETPEGRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSGAMAALA AEMBRE QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARRGGLGRYFGRGAA ASLVFVPTRRRGSPGSTAAMAALA AEMBRE QLPVSGVELDSQLLGKMAFEEQNSSSLH ASHGKDMVHKELHRPVSVSKLKYFFAVDTA VYAKKLGLLVPPYTHDWYQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVYSRDAPLPP RQDLNAPDLYIPTMAFITYVLAGMACH RCHAMATATATATATATATATATATATATATATATATATATA	]	1	1	1	1	1	
MPLELEWI,DVEHRITVAGVILSITHTGGVML	•	}				1	
LALVGYLIRDWRWLLLAVTLPCAPGILSLWW   VPESARWILTQGHVKEAHRYLLHCARLNGR    PVCEDSPSQEAVSKVAAGERVVRRPSYTDLF    RTPRLRHISLCCVVVWGGVINSYYGLSLDVS    GIGLINVYQTQLLFGAVELPSKILYVLSVRYA    GRRITQAGTLLGTALAFGTRILVSSDMKSWS    TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR    QTGMGLTALVGRLQGSLAPLAALLDGVWLS    LPKLTYGGIALLAAGTALLLPETRQAQLPETI    QDVERKSAPTSLQEEMPKQVQN    1051   2401   A 8759   515   1625   ERTPWAVSSAPSGDSEGDEETTQDEVSSHTS    EDGGVVKVEKELENTEQPVGGNEVVEHEV    TGNLNSDPLLELQCQCLCQLDCGSREQLIAHV    YQHTAAVVSAKSYMCPVCGRAEJSPGSLGR    HLIHISEDQRSNCAVCGARTISHATNSEKLP    EVLINESLPTVHNEGPSSAEGKDIAFSPPVYP    AGILLVCNNCAAYRKLLEAQTPSVRKWALRR    QNEPLEVRLQRLEBERTAKKSRRDNETPEERE    VRRMRDREAKRLQRMQETDEGRAERLQRDR    EAMPLKRANSTPEKRQARLIREREAKRLKR    LEKMDMMLRAQFGQDPSAMAALAAEMNFF    QLPVSGVELDSQLLGKMAFEEQNSSSLH    1052   2402   A 8763   1106   70   RIGHGGRDRRGGGRVARPGGLGRYPGRGAA    AGSKIFRARAFDPPPPPPDTDTSGGYSSQCGGY    PATGADVAFSVNFILLGDPMANVAMAYGSGS    AHGKDMVHKELHRFVSVSKLKVFFAVDTA    YVAKKLGLLVFPYTRIQNWEVQSRDAPLPP    RQDLNAPDLYIPTMAFITYVLLAGMALGIQK    RRSPEVLGLCASTALVWVVMEVLALLIGLY    ATVRSDLSTFIILLAYSGYKYVGMILSVLTGL    LPGSDGYYVALA WTSSALMYFIVRSLRTAAL   GPDSMGGPVPRQRLQLYTLGAAAFQPLIIY    WILTFILVR    RTPRUWYPELRELSAAAPRWSHRTAPGIMVF    YFTSSSVNSSASYTTYMGKDKYENGDLIKHGW    PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE    VLMDCAHLVKANSIQGCKMNNVNVYYTEW    SNLKKTADMDVGQIGHRRQKKRENEDLIKHGW    PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE    VLMDCAHLVKANSIQGCKMNNVNVYYTEW    SNLKKTADMDVGQIGHRACKRENEERKKREMDELRSY    SSLMKYENNMSNQDGNDSDEFM    1054   2404   A 8769   344   527   REATILACRNSCWYSRCSLGACKRTVCSMP		]		ŀ	,		
PPESARWILITQGHVKEAHRYLLHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVYA GRATITQAGTILLGTALGTRILLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALDGVULS LPKLTYGGIALLAAGTGRILLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALDGVULS LPKLTYGGIALLAAGTGRILLVBTRQAQLFETI QDVERKSAFTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVANSSAPSGOEDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNILNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAAVSASSYMCPVCGRALSSPGSLGR HILLHSEDQRSNCAVCGARFTSHATFNSEKLP EVNNMESLPTVHNEGPSSAGKDLAFSPPVYP AGILLVCNNCAAYRKLLBAQTTSVRKWALRR QNEPLEVVLQRLERERTAKKSRRDNETPERER EVRMMRDREAKRLQRGDTBGVARARLQRDR EAMRLKRANETPEKRQARLREREAKRLKRR LEKMDMMLRAQFGQDFSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRAA ASLVFVPTRRSGPSGGGRVARPGGLGRYPGRAG HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVHHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFYSVSKLKVFFAVDTA YVAKKLGLLVPPYTHONWEVQYSRDAPLPP RODLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVTVVNEHACHLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFGSBGYYVALAWTSALMYFIVRSLRTAAL GPDSMGGPVPRGRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENBDLIKHGW PEDIWFHVDKLSSALMYFIVRSLRTAAL GPDSMGGPVPRGRLQLYLHKGENBEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYYTPW SNLKKTADMDVGQIGFHRQKUKVTVTEKK VMEINRLEKTKVERFLHKERENBUPKE VLMDCAHLVKANSIQGCKMNVNVVVYTPW SNLKKTADMDVGQIGFHRQKUKVTVTEKK VMEINRLEKTKVERFPDLAAERGERDREER NISKKAQIQEMKREKEBMKKKREMDELRSY SSLMKVENNSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWFSRCSLGACKFTVCSMP		1	1	ł	ļ		
PVCEDSPSQEAVSKVAAGERVYRRPSYLDLF RTPRRIRISLCCVVVFGVDFSYPGLSLDVS GLGINVYQTQLLFGAVELPSKLLVYLSVRYA GRRITQAGTLLGTALAFGTRILLYSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAFTSLQEEBMFMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAFSGDSEGDEETTQDEVSRTHG EDGGVVKVEKLELNTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGF HLLHISEDQRSNCAVCGARTTSHATTNSEKLP EVLNMESLPTVHNEGPSSAEGKDIATSPPVYP AGILVCNNCAAPRILAGTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETTEERE VRRMDREAKRLQRMQETDEQRARRLQRMG EAMRLKRANETTEKRQARLIREREAKRLKRR LEKMDMMIRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLIGKMAFEEQNSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVFTRRSSGPSGTASVAAMAYHSGYGA HOSSHIFARAAPDPPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFYSVSKLKVFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTTMAFTTYVLLAGMALGIQK RFSFEVLGLCASTALUGLYTAAGHTLGLGYLA ATVRSDLSTFHILLAYSGYKYVOMILSVLTGL LFGSDGYYVALA WTSSALMYFIVRSLTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR  1053 2403 A 8768 2 712 RFPRVWYPELRELSAAAAFWSHRTAPGIMVF YFTSSSNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTEW SNLKKTADMDVGQIGFHRQKOKVIVTVEKK VPELNRLIERTKVERFPDLAAEEKGRNEER NEKKAQIQEMKRREKEBMKKKREMDELRSY SSLMKVENNMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKFVCSMP		1	}	1	1		
RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTALFTSELYFTVUR QTGMGLTALVGRIGGSLAPLAALLDGVULS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 ERTPYAVSSAPSGEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARTSHATTNISEKLP EVLNMESLPTVHNEGPSSAEGKDLAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDM EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMLRAQFODPSAMAALAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGAA ASLVEYPTRRSSPGSTASVAAMAYHSGYGA HGSKHRARAAPDPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRARAAPDPPPLEDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRAKAAPDPPPLEDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRAKAAPDPPPLEDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGYGA HGSKHRAKAAPDPPPLEDDTSGGYSQPGGY PATGADVAFSVNHLLGLAFKYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVSLETAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKONKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGERIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRKGDVKIVTVEKK VNEILNRLEKTKYERFPDLAAEKECRNEER NEKKAQIQEMKKREKEEMKKRREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSKCSLGACKFTVCSMP				Ì			
GIGLNYYOTQILFGAVELPSKILVYLSYNYA GRRITQAGTILGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYTTVIR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTOQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSNCAVCGARFTSHATHNISEKLP EVLNMESLPTVHNEGPSSAEGKDLAFSPPVYP AGILLVCNNCAAYRKLLEAQTFSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMGETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMLRAQFGQDPSAMAALAAEMNFF QLPVSCVELDSQLLGKMAFEQNISSLH  1052 2402 A 8763 1106 70 RHGHGGDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPLFDDTSGGYSSQFGGY PATGADVAFSVNHLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIYL ATVRSDLSTFHLLAYSGYKYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRAAL GPDSMGGPVRQRQLQLYLTLGAAARQPLIY WLIFTHJVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNIVNTVYTPW SRLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKYERFPDLAAEKGCRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 RRATTLACRNSCWVFSKCSLGACKFTVCSMP		İ	1	1	İ	ł	,
GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAGTALLPETRQAQLPETI QDVERKSAPTSLQLEEMPMKVQVN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGEEETTQDEVSSHTS EEDGGVVKVEKELLENTEQPYGGNEVUEHEV TGNLNSPPLLELQCPPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHSEDQRSNCAVCGARFTSHATFINSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLAQTPSVRKWALRR QNEPLEVRLQRHGERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREBEAKRLKRR LEKMDMMLRAGFQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRGGGRVARFGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTISGGYSSQFGGY PATGADVAFSVHELLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFCSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRUQLYLTLGAAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLESKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKPTVCSMP			1	1			
TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPVYP AGILLVCNNCANYKLILEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRNRDREAKRLQRNGETDEQRARLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH HOSS ASHGKDMVHKELHERVSVSKLKYFFAVDTA YVAKKLGLLVPTTRRNSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVYHLLGDPMANVAMAYGSSI ASHGKDMVHKELHERVSVSKLKYFFAVDTA YVAKKLGLLVPTHQNNEWCYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALA WTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKCKYCFMEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNIVYTPW SNLKKTADMDVGGIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNODGGNDDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKPTVCSMP	ļ	1		1			
OTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN	]				l		1 · · · · · · · · · · · · · · · · ·
LPKLTYGGIALLAGTALLLPETRQAQLPETI QDVERKSAPTALQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILHBEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEGRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMMMLRAQFQQDFSAMAALAAEMNFF QLPVSGVELOSULGKMAFEGQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGGAA ASLVFVFTRRSGPSGTASVAAMAYHSGYGA HGSKERRARAPDPPPLEDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRYVSVSKLKYFFAVDTA YVAKKLGLUFPYTHQNWEVQYSRDAPLPP RQDLNAPDLVIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVWEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYITLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKSCRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNDQGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		1	1	1	ļ	1	1
QDVERKSAPTSLQEEEMPMKQVQN					1		
1051 2401 A 8759 515 1625 EIRTPVAVSSAPSÖDSEGDEEETTÖDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLOCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLTPVHNEGPSSAEGKÜLAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRK WALRR QNEPLEVRLQRLERERTAKKSRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPTHQNIWEVQYSRDAPLPP RQDLNAPDTPTHQNIWEVQYSRDAPLPP RQDLNAPHTHATITYVLLAGMALGIQK RFSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTHILLAYSGYKYYOMILSVLTGL LFGSDGYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSNQDGSDEFM			i		1		l
EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPILELCQCPLCQLDCGSREQLIAHV YQHTAAVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNGGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANTEPKERQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTAŠVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVWMEVLALLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQXVLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAPRWSHRTAPGIMVF YFTSSSVNSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SINLKKTADMDVQQIGFRRQKDVKTVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1061	2401	<del> </del>	0750	616	1626	
TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILHISEDQRSNCAVCGARTSHATFINSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTPSVRKWALRR QNEPLEVRLQRILERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGYELDSQLLGKMAFEEQNSSSLH QLPVSGYELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTTHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1051	2401	A	8/39	212	1025	
1052 2402 A 8763 1106 70 RIGHGGRORANGENCAVCGRALSEPGELGR  1052 2402 A 8763 1106 70 RIGHGGRORANGENCAVCGRAPED STATEMENT QUESTERRE AWALKRANTEPER QUESTERRE AWALKRANTEPER QUESTERRE AWALKANALAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSLH  1052 2402 A 8763 1106 70 RIGHGGRORARGED GRAPP GRADA ASLVFYTRRSGPSGTASVAAMAYHSGYGA HGSKHRARA APDPPLFDDTSGGYSSQPGGY PATGAD VAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLLHKGENIEDIPKE VLMDCAHLVKANSIQGGKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFFDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1		1	1		1	
HILIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLMMESLPTVHNEGPSSAEGKDIAFSPPYYY AGILLVCNNCAAYRKILEAQTPSVRK WALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLATAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYVLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERPPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKEMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKPTVCSMP	1	l	ĺ	1	ľ		
EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQITPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGGDPSAMAALAAEMNFF QLPVSQVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVAPFGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLDDTSGGYSSQPGGY PATGAVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLIGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMMCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNELNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKKYENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ŀ	j			ŀ	1 3
AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRGGGRVARPGGLGRYPGGGAA ASLVFVPTRRRSGPSGTAŠVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKOMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			]	1			,
QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVOMILSVLTGL LFGSDGYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLBEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	İ	1		1	Į.	İ	
VRRMRDREÄKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENNSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1	1	ł	1	1		
EAMRLKRANETPEKRQÀRLIRÈREAKRLKRR LEKMDMMLRAQFGQDFSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTAŠVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		1	1		]		
LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	l						
QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYVYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1	1	1		1	1	
1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1		1	1		I	, , ,
ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			<u> </u>				
HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1052	2402	Α	8763	1106	70	· —
PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP						1	ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA
ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			1	1			HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY
ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP						1	PATGADVAFSVNHLLGDPMANVAMAYGSSI
YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1	1	1		Į.	1	
RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKEMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		1	1			1	
RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP						1	
ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		1	i		1	1	
LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP						1	
GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1		1		1	1	1
WLTFHLVR	1	1				1	1
1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			1	1		1	1 7 7
YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1052	2402	+-	0760	1	712	
PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1023	2403	A	0/08		/12	
VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			1	1	1	Į.	
SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			1	1		1	
VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			1	İ	1	}	
NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1		1	1		i	, ,
	1			!			
1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			1	1		Į.	1
						<u> </u>	
SLSRQGSQTLCLRLAEYCMESVDSQRLLLS	1054	2404	Α	8769	344	527	
		1		1			SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1 1100	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
'	]			peptide	-	/=possible nucleotide deletion, \=possible
		<u></u>		sequence		nucleotide insertion
1055	2405	A	87.70	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEK
ĺ		1				KMLKCVVVGDGAVGKTCLLMSYANDAFPEE
	1	1	1	1		YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ
				[		EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV
1		]	]		J	QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK
			1	1	1	TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL
		1	]	I	1	ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS
				1		EGHSCCSII
1056	2406	Α	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH
		1				RRDQKWHDKQYKKAHLGTALKANPFGGAS
,			İ		1	HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK
			1	-		NGKKITAFVPNDGCLNFIEENDEVLVAGFGR
ł		1	1			KGHAVGDIPOVRFKVVKVANVSLLALYKGK KERPRS
1050	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL
1058	2408	^	0000	11/1		VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME
1		1	1			TOSEPSELELDDVVITNPHIEAILENEDWIEDA
		1	İ		1	SGLMSHCIAILKICHTLTEKLVAMTMGSGAK
1		1 -		1.		MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL
						DPKLLDARTTALLLSVSHLVLVTRNACHLTG
				1	· ·	GLDWIDQSLSAAEEHLEVLREAALASEPDKG
		1		)	ļ	LPGPEGFLQEQSAI
1059	2409	A	8809	246	757	MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC
		1				EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY
1				1 "		LVWKDLGGGLGWPLALPLGLYAVQLTISWT
	1			ĺ		VLVLFFTVHNPGLALLHLLLLYGLVVSTALI
	1			1		WHPINKLAALLLLPYLAWLTVTSALTYHLWR
				1	<u> </u>	DSLCPVHQPQPTEKSD
1060	2410	A	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA
1	1	1	1	1	1	FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF
			-	1	1	GAMLNIAAVLCIATIYVRYKQVHALSPEENVI
			1			IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA
		1				HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH
		1		1	1	GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS
,				1		GNFGTDLEQKLHWNPEDKGYVLHMITTAAE
1		-	1	1		WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL
40.00	1	<b></b>	1000	<del> </del>	100	TLYDTAPCPINNERTRLLSRDI
1062	2412	Α	8824	) <sup>1</sup>	763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA
[			1	1		GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA
	1	1				YSLAPATPEVKVACSEDVDLPCTAPWDPQVP
		1	1		1	YTVSWVKLLEGGEERMETPQEDHLRGQHYH
1	1	1		}		QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET
1	1	1	1		1	FKKYRAEIVLLLALVIFYLTLIIFTCKFARLOSI
ļ					1	FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT
1						ELV
1063	2413	A	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE
1003	2413	1	0020	**′	1021	HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT
1	1	1		1		AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ
		1		,		KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH
	1	1				CVLSKEMKSVORSLGLSRIHLOSKRKIHFVL
}	1	1		}	1	TR
1064	2414	A	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE
1004	2714	1	0000	1 2702	1.005	LNKQVSELSQLYKEAQAELEDYRKRKSLEDV
		1	1			TAEYIHKAEHEKLMQLTNVSRAKAEDALSE
	)	}	}	} .	. }	MKSQYSKVLNELTQLKQLVDAQKENSVSITE
		[	1	1	1	HLQVITTLRTAAKEMEEKISNLKEHLASKEVE
		<del> </del>	·I			

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence	<u></u>	nucleotide insertion
	· ·				ł	VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
	<b>,</b>	ł			1	SLESEVSVLASKLKESVKEKEKVHSEVVQIRS
				}		EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
	ł	l		•		QAQEELAEMKRYSESSSKLEEDKDKKINEMS
		ĺ	1	{	ĺ	KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
		Ì		ļ		LQQQVKQLQNQLAECKKQHQEVISVYRMHL
		ļ	1	<b>!</b>	į	LYAVQGQMDEDVQKVLKQILTMCKNQSQK
	<u> </u>			ļ. <u>.</u>		K AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
1065	2415	Α	8841	3	663	
			1		ļ	APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL
	)	1	1	i	1	KDTTSSSSADATIMDIQVPTRAPDAVYTELQP
		i	1	1		TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
		1	1			DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA
		l			1	DEGITERSOFTEDDEET I DEGITERRAGELIA
	<u> </u>	1	100-	12006	2004	AVLFITGIILTSGKCRQLSRLCRNHCR FVGEQEGGCEAGAGRGAQTYPGEAGERWFG
1066	2416	A	8853	3806	2204	RRRRGRVVSRKKMSLKSERRGIHVDQSDLL
	1	1		ļ		CKKGCGYYGNPAWQGFCSKCWREEYHKAR
	1	Į.				QKQIQEDWELAERLQREEEEAFASSQSSQGA
		1			1	QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
		1		1		VGSKKEIOEAKAPSPSINROTSIETDRVSKEFIE
		i		1		FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE
	ļ			1		EQSECAQDFYHNVAERMQTRGKVPPERVEKI
•		1				MDOIEKYIMTRLYKYVFCPETTDDEKKDLAI
	1				1	OKRIRALRWYTPOMLCVPVNEDIPEVSDMVV
				1		KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI
						KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI
		1				QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
		1	Ì			KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
		1		1	1	WSPDACLGVKQMYKNLDLLSQLNERQERIM
						NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
	1	1				KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
1067	2417	A	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
1007	2417	1 ^	0000	13/2	1313	LRQAWATKQDPISKKK
1068	2418	A	8856	1530	1583	PCRPGMECNSMISVHCNL
1069	2419	$\frac{\Lambda}{\Lambda}$	8857	1530	1583	PCRPGMECNSMISVHCNL
1070	2420	$\frac{\Lambda}{\Lambda}$	8866	293	1675	PYPOGGYPOGPYPQEGYPQGPYPQGGYPQGP
1070	2720	1 "	0000	1	1	YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ
1	1	1			1	EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
ľ		1		1		LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
!		1				WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
		1	ł	1		VALSVLTASLSYMVGMIASFYNTEAVIMAVG
1			Į	1		ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
l		1		1		VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
		1	1	1		VDTOLLLGNKQLSLSPEEYVFAALNLYTDIINI
ļ	1	ł		1		FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
				1		WHGSASCTSPLSCPQAQPREKDASLQPSCMY
	1	1	1			TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
1	1	1	1			HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
		1		1	1	EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
ŀ		1				GDMRSGGLIPVLSPE
1071	2421	A	8868	+	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
10/1	2421	1^	0000	1	1 330	DDKMGSNTFFKRNDCRYVMISCKADMAYDN
I .	1	{	{	1	1	VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII
i		1	Ì	1	1	LNGEKLKYFPVRSGT*QGCSVWP
	ļ				1	
1020	2422	1	2070	33	658	MESVLSKYEDOLTIFTDYLEEYPDTDELVWII.
1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKOHLLKTEKSKLLSDISARLWFTYRRKFSPI
1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH

SEQ ID NO: of nucl- cotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						CYSIHQMAQMGVGEGKSIGEWVLGPNTV\AQ GV*KNLA\LFDEW\NSLGLVYVSM\DNPSGSIA RFPKKLCRVLPL\SADTAGLTGP
1073	2423	A	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL  *QVCITA*IKESVIISGG*SSSPVCHTTFQPANL RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW KEISFGDYICHTFQGDCWADRSPLHEAAAHG RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL *GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	A	8896	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK PWPSLLDKEREESLRQKRLSERERIGELGAPE VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS TSEEEKKKKSSRSKERSKKRKKKSSKRKHK
					,	KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK EKKKKHRSKKYKKKRSKKSRKESSDSSSKES QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGEGAAMAEYVKAGKRI PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR MEAVRTAKREPESTVLMRREPLHPFNPRRET KERE
1076	2426	A	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E *APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW QYPALHRAGTEWQLSALHRAPRSTQPDKAC RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT \YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	A	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRRQHLHGPLH RP*QAAPALHAGCQLAPHPPT
1079	2429	A	8912	121	376 -	NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY
1080	2430	A	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC EERTKMSTGPDYKATVGDISSDGNLNVAOEE
1081	2431	A	8922	56	420	CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVP\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST
1082	2432	Α	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

**************************************	CEO ID	1 3 4 - 4	1 880	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	peptide	liou.	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	nence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	целее	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciicc		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
	Ì	1		sequence	1	nucleotide insertion
		<del> </del>	<b></b>	<del>                                     </del>		GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
			İ	ļ	1	QGTQIASDGLKGLLFEVSLADLQNDEVAFRK
	1	ł		İ	ļ	FKLITEDVQDKNCLTNFYGMDLTCDKICSMV
	ŀ	Į.				EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK
						HNNQILKTSYA*HQQS/RQIQKKMMBIMT*EV
		}	ļ	1	ì	QTNDLKEVVNKLIPDNIGKDTEKV/CPIYPLH
				l		DVFIRKVKMLENPGFER\MELRGGGSSS
1083	2433	A	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP
	ļ	1	1	*		WGAVQGSRAMSDLLLLLLDLTLLLLLMLLGF
	İ					AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
		ł	}	l	<u> </u>	GWLLT/ESCSISPKLCSIAVH*DNPAWF
1084	2434	A	8950	156	318	HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
		-		<u> </u>		WGGKRVQPFWKRVWQKRTLNLRV
1085	2435	A	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
	İ	1			İ	*TIYTSYDTAIPIS/GI/YPKRMSSKCHQETCAR
	ļ	į				MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
			1			HKYYIKVKRNL*VTITH\TWVNLNILMFEIILW
			<u> </u>		1006	YSHKYY H*KILQVGRAQRAHXSRL*SQLLRRLRHESHL
1086	2436	A	8962	868	1026	
	<u> </u>	ļ		<del> </del>	220	NPGARGCSEARLHRCTPAWTT LHVKHLGHFQLVFSEVICHCILMPVS*ELQRL
1087	2437	A	8985	58	330	
						*ERSVCAFHVCIQTYVCLQVYACMCVYYICM FVYSVYGCGLCTCVCMDVYICVCVQEFL
		<b>_</b>	1	<del> </del>	404	N*KWILHVNVRIQSIFF/KRNQK/INSHELKLD
1088	2438	A	8989	394	404	KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA
						KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV
						YKDLSKLNRRKTE/S*/VKKWVKDLSRYFIKE
i		ŀ		ł	Í	VISMENKHKKIFSTS
1089	2439	A	8991	60	329	MALTPESPSSFPGLAATGSSVPEPPGGPNATL
1009	2439	^	6971	00	327	NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV
		1		)		GVEDNAYTLEVNSRYMRAVGIM*IHL
1090	2440	A	8996	2	351	SNITITLT*MKKYDNTFCW*GCGQIG/T/LIYC
1050	2770	1	0,,,0	-		WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL
				j		GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR
1	[	[	1			LLIAALFIIVQYWKQSKDHYI
1091	2441	A	8997	97	456	YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT
						LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG
-	ì					AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS
}		1	1	,		FDKNYKAPIGADFEMERFEVLGIPF
1092	2442	A	8999	548	811	SSFIKRHILIFEDDWHQTTCCHHPHHP\F*RCQ
						FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH
		]	]		1	RAAHHHQHGQGPLGHGLVARVG
1093	2443	A	9002	3	2745	ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS
						TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD
		1	1	1		AMCCLRYWYTPESWICGGQWREYFSALRDF
			İ	ļ		VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR
	1	1	1	1	1	LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA
				1		VFTRFALKTLGQETLCSLQEADYEVASYGLQ
		-				HNCLGILGEDTDYLIYDTCPYFSISELCLESLD
		)		1		TVMLCREKLCESLGLCVADLPLLACLLGNDII
1	1				1	PEGMFESFRYKCLSSYTSVKENFDKKGNIILA
		1	1	1		VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL
	1		1	1		*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP
		1		1		RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP
ļ	1		1	1		GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG
į.	1	1	1	1	<b>S</b>	PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP
l l		- 1		1	1 .	ACTCDE AD OBVDNCTOCEDD OBVDNCTOCED
ļ						MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP
						MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY TGPESROEVLIRTDPESROEIMCTGHESKQEV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						PICTDPISKQEDSMCTHAEINQKLPVATDFEFK LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS DTEILKVARTHHVQAESYLVYNIMSSGEIECS NTLEDELDQALPSQAFIYRPIRQRVYSLLLED CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA CFNLSSSREELQAVESPFQALCCLLIYLFVQV DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP DYINPRAVQLGSLLVRGLTTLVLVNSACGFP WKTSDFMPWNVFDGKLFHQKYLQSEKGYA VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV
1094	2444	A.	9021	97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE T*HIMAEPVSPLKHFVLAKKAITAIFDQLLEFV TEGSHFVEATYKNPELDRIATEDDLVEMQGY KDKLSUGEVLSRRHMKVAFFGRTSSGKSSVI NAMLWDKVLPSGIGHITNCFLSVEGTDGDKA YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRVFWPKAKCALLRDDLVLVDGPGTD VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	A	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW LLHRRARRSSALCPRPRSWGVSGGEGAGARE P*ITSSSCCLSAA/SHLSIQSPNMAGARRRIRPQ LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV FAYGQT\GAGKTYTMGTGFD
1096	2446	A	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRES*SCPYTPAWITERDPV
1097	2447	A	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	A	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK
1099	2449	A	9043	185	372	IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI
1100	2450	A	9045	763	584	RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF
1101	2451	A	9050	275	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF I*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA AGMQI/H/CWWCVNVGKFWEMS*YYLLKLSI ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

OFO ID	L CEO III	1 3 4 44	CEO	Th	15.0.0	· · · · · · · · · · · · · · · · · · ·
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	{	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	l	1	peptide	1 -	/=possible nucleotide deletion, \=possible
			1	sequence		nucleotide insertion
	<del>                                     </del>	<del></del>	<del> </del>		<del> </del>	APFVLAVNC
1104	2454	A	9064	75	393	
1104	2454	Ι Δ	3004	1 13	373	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI
	1		1			KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V
		1	1	İ	ļ.	KTDCGCGANSKGVVVVMKV\KTAQQKQTTS
			<u> </u>		<u> </u>	YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS
			1		į	RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR
	1		i		į	AWPCCPGWSAAWLTIVILAHYRRPGLERSCC
		ļ				LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL
	1		ł		1	VLNS*TQGI
1106	2456	A	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT
	1 - 100	'-		0,0	0.0	HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	PROCEED ATVICTOR ATTERNATIONAL TRANSPORT
1101	2431	Α.	7000	1 200	۰، ا	KPSSGSFIRATYIFLSTAHVPALFSVLVRTKLT*
			1	1	l	AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA
	1	]	1			VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL
	1	ł				I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA
		1			1	PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP
	l	<u> </u>	<u> </u>	İ	1 .	ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	A	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP
				ļ		GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA
	}	1				SAFPPAERSRGHRRASL*RARWSAAVPRRSA
	ì	l	Ì	İ	ì	GSASEPVOSRWLRLPVGSDSPPAVPVRVCPAP
			1			DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG
	l	l	1			1
1109	2459	A	9099	1255	1406	QRPPPPSGDSLSPPGCCRY
1109	2439	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL
1110	2460	<u> </u>	24.55	ļ		GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	A	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT
		J	ļ	1		CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP
	i	l	ľ			SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS
	l			ł		LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	Α	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP
	į					AAAGDPASLDFAQCLGYYGYSKFGNNNNYM
						NMAEANNAFFAASEOTFHTPSLGDEEFEIPPIT
	1	i				PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT
	}	l	ļ		}	
						PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM
					Ι ,	DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG
		l	1			VMPPAQLTTINQSQLSAQLGLNLGGASMPHT
	1					SPSPPASKSATPSPSSSINEEDADEANRAIGEK
	1	<u> </u>				RAAPDSGKKPKTPKK
1113	2463	Α	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F
		1				SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH
			i			VGQAGHEPLTSGDPPASASQSAGITGVSHOA
	1	1	1			WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS
						LLKCWDY
1114	2464	A	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS
		1	1			SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII
		Į	1			
1116	2465		10104	<del></del>	001	YLYVQTVPQRV
1115	2465	Α	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG
	1		1		İ	SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM
	1	ł	l			TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE
		Į.		1	1	WRLQHKGRGRGDLHLPDHHLSVPSSADHPA
						QQPSQFNGRNLYFLPLFR
1117			0125	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA
1116	2466	A	9135			
1110	2466	A	9133			OHVHVPPWTDVI AGODDD A DT AGOG A DIVID
1110	2466	A	9133	,		QHVHVPPWTDVLAGQDRRAPTAGDGAPWP
1110	2466	A	9133	,		QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ
1110	2466	A	9133			QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA
1110	2466	A	9133			QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTITIELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG* GILTKLSAFLTIPRLQPHLIAALSPSS
1118	2468	A	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	124	207	ACPRLARRRRVRSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLEKECLEK ESRDYDVDHPGEADSV/LRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFFNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLPWQRRGLLRAQGLRG\ WKARRAITTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEDEDDEDGGEEAPAPG GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFPKKISSKSLKVGMIPAPKRVCLIQE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1121	2471	A	9166	272	523	PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
1122	2472	C	9170	442	236	V/CSHITDSLKFIGKGWVGMVTHACNPGTLG G*GGWIA*VREFETSLGNM MNRRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN* MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC

070 70	OTO TO	Max	CEC	Dandies	Deadist-11	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	denoc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ		[	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l	1	peptide	_	/=possible nucleotide deletion, \=possible
				sequence	L	nucleotide insertion
1124	2474	A	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
			1			WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
	]		1			TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA
	2455		10150		100	IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1125	2475	A	9179	704	188	ESSGLLFQCFQGHVQKLTLQARPTLFSWWL
		ł			Ì	CSKPPKETGELENAESGGDGGRRGGKQDNV AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA
		1				LPMGFFYLYFRDPGREITWKHFVQYYLARGL
		1		1	· ·	VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE
	[	1			•	YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	Ā	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM
		} '-	1	-	}	EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
		1	1			ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
ļ t	ı	j				LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL
	ľ				ļ	RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
	!		}			L\LDGVPVALKKVQIFDLMDAKARADCIKEID
	İ		ĺ	İ	ĺ	LLKQLNHPNVIKYYASFIEDNELNIVLELADA
			1		1	GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS
	ł	1	İ			ALEHMHSRRVMHRDIKPANVFITATGVVKLG
1	}	1	1	ļ	1	DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD
1100	2479	<del> </del>	9190	1	370	NG GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
1129	2479	Α	9190	1	370	PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
			ì	İ	1	RATTIKIRVVATITRARIEDMRHSATALTRPD
		1	1			ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
	1 -100		1	"" "		LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN
1	1	1	1	ł	ŀ	PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
			1			DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE
		J		]		PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
	1	1		-	Į	CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
					Í	GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
	0.400	<del>                _  </del>	10006	<del>                                     </del>	950	AATGVYIFFQNKY
1132	2482	A	9206	] 1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK TQPVEATDDAFWDQFWADTATSVQDVFALV
				}		PAAEIRAVREESPSNLATLCYKAVEKLVQGA
				1	Ì	ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
]			1	1		WRGFFWSTVPGAGRGGQGEEDDEHARPLAE
1		1		1		SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH
1		1		1		SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME
1		1		1	1	LLKLLLTCFSEAMYLPPAPESWQH/RTHWFSS
}		1	1	Ì		FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
				1		NHLY
1133	2483	Α	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
ļ	1	1			1	AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
-				1	1	HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS
				<u> </u>		NVYFIV
1134	2484	A	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK
1 .	1	1				RADGAAPAGEGEGVTLQGNITLLKGVAVIVV
1	1				1	AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
1	1	1			1	GVFSIVGALCYAELGTTISKSGGDYAYMLDV
	}	1		)	}	YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
		1	1			YSVKAATRVODAFAAAKLLALALIILLGFVQI
	1	1			-	GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG
	1	}				LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP
	_L				<u> </u>	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			ì			IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF
	[					GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS
			1	{		RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT
		ŀ	1	}		CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII
		1	]	<u> </u>		GMIWLRHRKPELERPIKVNLALPVFFILACLF
				1		LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV
	<u> </u>	<u> </u>			100	WWKNKPKWAPPGHLSPRPSCVRSSCMVVPQ
1135	2485	Α	9216	40	410	RDRLPPAYFCRPVVCVVTALDVG\SPESQEM
	ļ			1		DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV MQETLRNLASIGEKWKDQNIEDQYKNPRNNL
		1.				
		<del>  </del>	1	<u> </u>	002	RSLLGERVDENTEENHCGETSSQIPDDTLNK RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG
1136	2486	Α	9223	3	983	DLVFAKMKGYPHWPARIDDIADGAVKPPPN
	ì	Į.	1			KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGK
		1	i	1		PNKRKGFNEGLWEIONNPHASYSAPPPVSSSD
			1	ł		SEAPEANPADGSDADEDDEG\RGVMAVTAVT
			1	i		ATAASDRMESDSDSDKSSDNSGLKRKTPALK
			1	l l		MSVSKRARKASSDLDQASVSPSEEENSESSSE
		1	1		1	SEKTSDQDFTPEKKAAVRAPRRGPLGGRKKK
		1	1 .	1	-	APSASDSDSKADSDGAKPEPVAMARSASSSSS
		1		l		SSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK
•		1		}		PKPERPPSSSSD
1137	2487	A	9229	21	239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV
1137	2407	^	7227	1	233	GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL
			1			RGSREPPAWA
1138	2488	A	9231	1664	2	TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL
1130	2700	1	720.		-	EGIVWHETEEGVLVVNVTWRNKTYVGTLLD
	1		İ			CTKHDWAPPRFCESPTSDLEMRGGRGRGKR
		1	1	i	ŧ	ARSAAAAPGSEASFTESRGLQNKNRGGANGK
ĺ	1	1				GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR
	1	1				KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP
ļ		1		1		TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK
		1		ŀ		YKHINGLRYHQAHAHLDPENKLEFEPDSEDK
!				1	1	
	1		1	F		ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA
1						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA
					_	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKN\PSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKN\PSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPLAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS
1139	2489	A	9234	207	443	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS
1139	2489	A	9234	207	443	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL
1140	2490	A	9238	248	328	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRAAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV
						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP
1140	2490	A	9238	248	328	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG
1140	2490	A	9238	248	328	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA
1140	2490	A	9238	248	328	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNIPSKLKSARPILAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRKALAPPLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\
1140	2490	A	9238	248	328	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPILAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGMN WPWAAALVVHCYSKSPSNKDAALLEAARAQ
1140	2490 2491	A	9238 9242	248	328 535	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPILAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLLVATGG RPGTSPALFSGRGAATGGRQGGFPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGMN WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN
1140	2490	A	9238	248	328	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPILAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD
1140	2490 2491	A	9238 9242	248	328 535	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKIKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \UMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC
1140	2490 2491	A	9238 9242	248	328 535	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGMN WPWAAALVYHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN HICFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF
1140	2490 2491 2492	A	9238 9242 9245	248 2	328 535 466	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ
1140	2490 2491	A	9238 9242	248	328 535	SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGMN WPWAAALVYHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN HICFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	nence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1	i	peptide	ĺ	/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP
		}	Į.	ļ	ŀ	SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE
	ł		1	Į.		AMEESDRPCEISEIDDNPKISENPRRSPTHEKN
	ŀ	Ì				TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM
		ļ.,_	0024	100	411	ERRR METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF
1145	2495	A	9264	175	411	PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL
	1	1			1	WDTAGOERFISIT
	1-100	<del>  </del>	0000	500	814	MFTYLEGREGIKSQPKMEPHSVT\RLECSGMI
1146	2496	A	9277	592	1 014	SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA
	1	ļ		i		WLIFAFLVETGF
1147	2497	A	9279	1255	2	FRRGRRGEEEKEEEEEEEGWVNGMENSHPP
1147	2491	10	7417	1233	1"	HHHHQQPPPQPGPSGERRNHHWRSYKLMIDP
		1	1			ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE
	1			ļ		DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV
	1			1	1	PPKOVTFAKLNDNIRENFLRDMCKKYGEVEE
						VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ
1	1		}	}	1	HLHSTSVMGNIIHVELDTKGETRMRFYELULV
-				1		TGRYTPQTLPVGELDAVSPIVNETLQLSDALK
	<b>.</b>	1				RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA
1	1		1		1	YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY
ļ	ł	1	1	Ī	1	SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY
İ		1				NRRHEHHYVHNSPAVTAVAGATAAFRGSSD
ļ	1	l		1	1	LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP
	1	ļ				PAQATPAPGFR
1148	2498	Α	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY
		1		1		ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL
Į			ļ			SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN
			Í	[		LSLPDV\GFTSTTVPK\MIVDI\QSR\$RVI\$YAG
			1			CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI
	ļ		1			CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL
1	į.	}	1	ł		HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD
1			1			SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL
İ		1	1	ļ		GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT
1	1			1		STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF
11/2	10400	+		1	600	GGMEERHAPECDGL MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE
1149	2499	A	9303	1	699	FRLVAADRSMGRYMLFGVINLICTGFLLMWC
ļ			1			SSTNSIALT\SYTYLTIFDLFSLMTCLISYWVTL
1	1	1				RKPSPVYSFGFERLEVLAVFASTVLAQLGALF
}	1	1	1	1	]	ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF
	1	1				TMLSIRNKPFAYVSEAASTSWLQEHVADLSR
	1			1		SLCGIIPGLSSIFLPRMNPFVLIDLAGAFALCIT
						YMLIEI
1150	2500	+	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL
1130	2500	1^	3300	1'3'	""	SLP
1151	2501	A	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG
1131	2301	1 ^	7309	203	1 ***	SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ
	-	1	1	1		FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV
1132	2302	^	7517	1	1	OGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR
1	1	1	1	1	[	PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I
	-	-	1		ĺ	PPTSA/FPKCWEYRSEPOCLPGCLSFSGILLDL
	-	1			1	GTNVSLRAA
1153	2503	I A	9315	392	<del>                                     </del>	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP
1133	2303	1^	1313	1 372	1.	PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG
1		1				GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA
1			1 .			PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR
1	1	-	ı	1		PGNS
1	1					

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	ł	}	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
	L	<u> </u>		100		PT
1155	2505	Α	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Α.	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
(		ł			1	EAAQRHCRASVSILRMRRPGQGSSRPARVPL
11.65	0500	ļ <u></u>	0007	150	1000	RGPDSHRLREPPPSPP
1157	2507	Α	9327	152	292	YERRGRSQGGGSHPAGAQPGGRAIGAGWQS KEPLWEGLQRSGSPLPG
1150	2508	<del> </del>	9328	<del>                                     </del>	430	QELKOGPNPLAPSPSAPSTSAGLGDCNHRVD
1158	2508	A	9328	1	430	LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
	i	]	į.	İ		VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
	-		1			SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
		1			}	RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
1137	2307	n	7554	100	303	LMEAGLPQKQAERADELFEAGLVIYVKLDER
			1			VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
1.00	2570	1		-		KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
						DHTDQELREEIHKANVERVVHDVSQEATIEKI
		1	1		}	RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
		1	1	}		EAELPIMSQLTEIETCVEC
1161	2511	A	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA
						AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
				1		DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
ļ					1	QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
}	ŀ		1	1		GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
					1	TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
1						VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
		Í		1	1	VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
						GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
						FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
}		1	1			WIDGPKDTSVEDALERTYVPCLKTLQEEVME
				1		AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
	10510		0246	0.0	212	G DOY AL OPPLY COCCATE A LICHT TRIPCOTTERS OF G
1163	2513	Α.	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
	(	1			1	LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
Ì		}	l			FSSYQRNNPDLILNDTIMPNIK
1164	2514	1	0347	13	1000	
1164	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
		1	1		1	FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
		1	-	1.	1	VAGASVGAGVWARNPRYRTEGEACVEFKA
1		1	1	1		MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
1		1	1	1	1	VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
						SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
			1	1	Į	SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH
		1	1		1	VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS
	1	İ	1	1		YVSIFVPLWLSLLTLMATTFRRKGGNHWWF
			1	1		AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA
		1				LPLQNKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
****	122.5	1.	1	1-"	1	VPEGVRLADGPGHCKGRVEVKHONOWYTV
				1		CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
1	1	1		1		TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
ł	1	1	1	1	1	LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW
		<u> </u>				PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168	2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS QQSILAGLVVVATTGMIGSPLECLFGELGGRA DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN NSARRMEAMASGSNWLSGVNVVLVMAYWS LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH NAPKDLKEEIDILLSRVHNIKYEPHLLADDDA
1170	2520		9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF TNETWQARTGEPLPDHLVLLMWSLIVSLYPL GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS AAILFGFSRKAGSFEMIMLGRLASWGVNAGV SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA LGIVMGQVVGLSTTAATGLRGL\AGELEELEE ERAACQGCRARRPWELFQHRALRRQVTSLV VLGSAMELCGNDSVYAYASSVFRKAGVPEA KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL WGGTPRSFALNQFTLQKKKK
1171	2521	A	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG SKRLFFWMQEP
1172	2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKET\KAAMGTQ CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL ASFNEVGNTALIVLESY
1173	2523	A	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI* KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV IRPPISFSKINNGP
1174	2524	A	9397	77 _	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF LGML
1175	2525	A	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT AKMNKWSKGKVRDKLNNLVLFDTATYDKL CKEVPNYKLITLAVVSERLKIPGSLARAALHE LLSRGLI*LVIQHIAQVIY
1176	2526	A	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY GERGYAQNGDF*DAQLDDYSFSCYSHAQVN GAPNSLTRAYDDP*VKISGLECQKVGALVEV KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1178	2528	A	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA WWRAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1180	2530	A	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK L
1181	2531	A	9436	2	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP IARTILDRLTGIPHGYCFVE*ADWATADKCVH IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL SMILK*MGAGDEKISAMGKARVDHRELYLGL LYPTEDYKLTFRARH
1183	2533	A	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR KGCSGWAPWLSLQCQHFGRPRWADHLRSGV RDQPGQYSKTTFLPKIQKLAGHSGAHL+S+LL ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT ERG
1184	2534	A	9462	391	655	LSGFKSLMPKIPLQYIYVRVRTTWSFCLPLDG RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	A	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA WWWGWECWVRALKLSSGPAGPLACWVAK KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG WSAVVQFWLTAASNSCFSLLSSWDYRCA
1186	2536	A	9468	275	452	HIPQLHTKTHYVPTRMVNKI*QIDNSKPWQR GG*TGILTHCW*ESKLVQPLWKIVWHYQ
1187	2537	A	9469	388	3	EVAPGPSQILPRRVTDGGDRPOFSLPGPRLPQ SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN S
1188	2538	A	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI
1189	2539	A	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	A ;	9489	i	411	LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK
1192	2542	A	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS
1193	2543	A	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	A	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP GTRTPSGCQNPAASGG

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ		1	peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
1196	2546	A	9518	sequence 229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
1190	2540	^	3310	223	700	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
	•	Ì	ł		į	GYIGALFPMSGGWPGGQ
1197	2547	A	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
						HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE
			<u> </u>			APNWKYKYGY*IPVDMLC
1198	2548	A	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
1		ĺ				VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK
1100	2540	<u> </u>	9546	1785	1943	SSYS GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
1199	2549	A	9346	1785	1943	V*QRGDGKNPGVTHLNRPVGTX
1200	2550	A	9548	186	1	VNAEKEP*KIQHYFMTKSQNKLHIEHTYLKPI
1200	2550	[ ^	7540	100	1	KAIYDKWTSDIMLNLOKL*AFFLRVIVROI
1201	2551	A	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
		1				GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSQ
i		1				YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS
ł		l		}	1	KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
<b>!</b> :			İ			PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
1	ļ	1		j		PDVHFFHCDEVEAELVHEYMESALTDCRLGK AMRP
1202	2552	A	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
1202	2332	^	7552	420	*	LDCERPPQGPLPSLPELAKTSYSDLTGLATED
						*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
	į		1			LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
						SKPRATPPLFCSLHTF
1203	2553	A	9568	517	738	RRKFERKQKQ*RYREGKQYRQRDKMKEWG
	i				] .	EKEKRREKGEREERKMRHRERKGESGQRD
1204	2664		9573	83	415	TMENWRVERLTEKER EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
1204	2554	Α	95/3	83	415	DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
	1	{	İ		į	EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ
						HDCRHKEDAGVICSEFTALR
1205	2555	. A	9577	64	424	ARGSCPTRPRTANGRMGETKDAPQMLVTFK
						DVAVTFFREEWRQLVLVHRTLYR*GMLETC
		1		1	`	GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
1006	10000	<del> </del>	0504		L	VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
			1	1	· ·	NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
1		1	ł	ì		GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
			1			YNPGLPPLRTWNGQKLLWL
1207	2557	A	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
1			1	1	1	PDGCRNVLRPKYYRLCDKAESWGIALETVPT
		1.				GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
1		1	1		1	THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
1208	2558	1	9597	122	3	FGILFSICFS IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
1200	٥ددع	A	19391	122	"	FADAWADAW
1209	2559	A	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
-207	2007	1	1	1	1	GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
	-	1	1			RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
		1				MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
				<u> </u>	<u> </u>	LLNASITETFNC
1210	2560	Α	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
		1		1		DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
		}			1	KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS
		1		1		GELSAK
1211	2561	A	9620	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR
						/

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghtamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
						GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	Ā	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST EENFGEKLHDIGFGNGFLDKT*KAQATKAKI DK
1216	2566	Α .	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI NHLPETERNLLEHGLMYIRLNAAFCSLVAHS LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV EEHHLQPVQVLQTLLHSATAGTGCRRPARPP PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL GALGGRGGRALGGSRWPPPLPGETLFSGCKH RRRRGSDAAPGEEAGT
1218	2568	A	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF MLDFEGEDTFHGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL
1220	2570	A _	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K
1221	2571	Α .	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS
1222	2572	A	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	A	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGHVRIVSFL LKKECLCQPQKPERENLLALCCE
1224	2574	A	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE GRRPRAVKVYTINLESQYLLIQGVPAVGVMK ELVERFALYGAIEQYNALDEYPAEDFTEVYLI KFMNLQSARTAKRKMDEQSFFGGLLHVCYA PEFETVEETRKKLQMRKAYVVKTTENKDHY VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	1	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	[	09/496 914	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		{	914	ng to first amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide	Sequence	/=possible nucleotide deletion, \=possible
		1	ł	sequence		nucleotide insertion
1225	2575	A	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
1223	23/3	^	7710	1 *	105	TDDVHINDISLQDYIAGVRLILL
1226	2576	A	9713	82	492	QGLPSFLPAFQPSGSWLGPAPTLGSSCNTVDT
1220	2370	, °	'' ''	42	""	ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
		l	l	1		ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
Ì		· .				WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
l	1	ł				MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
•==-	]		****	-		LAAGKMNISIDLDTNYAELVLNVGRVTLGEN
		l			ļ	NRKKMKDCQLRKQQNENVSRAVCALLNSGG
Ì		i		1	[	GVIKAEVENKGYSYKKDGIGLDLENSFSNML
ļ						PFVPNFLDFMQNGNYF
1228	2578	A	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
		1				TLVLQKSDVEAVF
1229	2579	A	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY
İ		1				GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
	ļ	1			1	QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
						PLLEELGINFDHIWQKTLTVLHPLKVADGSIM
	İ	1	İ			NETDLAGPMVFCLAFGATLLLAGKIQFGYVY
	i	1	1			GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
	, ·	1				GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
	1					WCSFSASKIFISALAMEGQQLLVAYPCALLYG
		1		·		VFALISVF
1230	2580	Α	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
i		ł		}		HFSPERPFMDYFDGVLMFVDISGKCKRDVCL
				l	<u> </u>	MWMSNRLAWEFTCRA
1231	2581	Α	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
1		1	1		]	ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
		1		1		LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
	1	}	1		}	LRCGWSPAEELNYTVPGPGPAGEASPRQCRR
		Ì			į.	YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
	1				j	RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI
1		1				NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG
			1.		1	WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL
		i				LVLAGVAYALPHWRWLOFTVALPNFFFLLY
ļ		1		1		YWCIPESPRWLISONKNAEAMRIIKHIAKKNG
			1	1		KSLPASL
1232	2582	A	9753	164	517	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC
1432	2002	1	7,33	107	1 ""	YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG
		1	1	1		LLOTVFVSCLLLSAPVFGYLGDRHSRKATMS
			1		1	FGILLWSGAGLSSSFISPRYSWLF
1233	2583	A	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
1	روري	1.,	1	1	1	IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
1		1	1		1	KRAYKSYVRALPLLKKMGINSILLRKSIGALE
]		]		1	Ì	VACGIVMTLVPGRPKDVANFFLLLLVLAVLF
				1		FHOLV
1234	2584	+A	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP
*~~*	2504	1	1	1	1	CSRCGYGVYPAEKISCIDQIWHKACFHCEVC
1	j	1			1	KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS
	1	1	1			VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
	1	1	1			FHWD
1235	2585	A	9767	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA
		1	1	1	1	SPODLSSGLAVAYVLNQIDPSWFNEAWLQGI
			1		-	SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP
						ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
	1	[				CARPGHSARNNTDKNLPHTAILLVTSNTYTTI
						KINFQAGRSGSCL
1236	2586	A	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC
1 1236	2586	I A	9/70	332	1 008	TRUEALI VATLIARTIGETASNEESITAAHLC

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ì			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	)	}	İ	peptide	j	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	1					LERKQLNLEIYDPCSQTQKAKFSLTSELHWA
	2505	ļ	0700	266	616	DGFVIVYDISDRSSFAFAKALI
1237	2587	Α	9793	266	515	NILAIIYFPFPRLFLLRDSQSNPKAFALTLCHH
	Ì	}	1			QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS   RYKPTRPVCITQFQGCS
1238	2588	A	9802	537	967	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
1430	2300	^	7002	337	, 307	FRDCNKIAFYRRQKQWLSKKSTYRALLDSVT
		1		ĺ		TDEDSTRFOINEASKVPLLAEIYGIEGNIFRLK
	1			İ	ŀ	INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
		ŀ		1		GSLILADGKGDLKC
1239	2589	A	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
1277	2507		1000	1.00	• • •	MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
		1	1	ļ	,	GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
	į			1		PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
		ſ				DIKFLTKGDNNEGDDRGSYK
1240	2590	A	9819	3	305	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
						PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
	1	1	Ì			YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF
					,	DPNSEEHG
1241	2591	A	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
	ŀ					LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
	•		}	j		GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
						SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	A	9843	3	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
						LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
	1	İ	1			DGSVKCTHQPCLIKWISERGCWSCELCYYKY
ł	1	1	1			HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
		İ				LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
						MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP
1243	2593	A	9846	198	411	WRISHHAGKMPVMKGLLAPONTFLDTIATRF
1243	2393	1.	7040	170	7	DGTHSNFILANAQVAKGFPIVYCSDGFCELAG
ļ		1			ļ	FARTEVMO
1244	2594	A	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
	237.	1	70.0	***	1 323	PANGAGGPGGASARKLSTFLGVVVPTVLSMF
ſ	1	i	1	ĺ		SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA
	1			ŀ	İ	LTVLSVCAIATNGAVQGGGAYCILQHRWTG
	İ				,	VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
1		1		1		NVCGCAVSLLGLVESVLDVFGA
1245	2595	A	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE
1	1	1		1		AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
]	Į.	}			}	SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ
		1	1			AMSLTEDVLAAALADHLPEDKWSAEKRRPL
	1	1		İ	1	KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
1	1		1	ĺ		YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
		1		1	1	RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
	1	1	1	1	1	SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT
		1			1	NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
1		1		1		EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
		1	1	1	1	LSGPTSEGLMTWELDRLLWARSVENLATATT
L		<u> </u>	100==	<del> </del>	<del> </del>	TLTSLA
1246	2596	A	9850	114	464	PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS
						ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
1				1		QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
10:5	1	<del>  </del>	10051	<del> </del>	1 205	RRERPAGADSLSWGAGPRISSYV
1247	2597	A	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
	1	1	}	1		HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP
1 '	1	1	ì	1	ì	GAILCSKHFQESDFESYGIRRKLKKGAVPSVS
1		1	1	ſ	ľ	LYKVFKYSSRCTS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RVDDFVYSKGGKDAGGADVSLACRRQSIPEE
1240	2550					FRGITVVELIKKEGSTLGLTISGGTDKDGKPR VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR LRHDEITTLLKNVGERVVLEVEYELPPPGGCP WT
1249	2599	A	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA SGSGVAAGPAARHAPRRRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTTLLALLALVLLYLVSGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVTLTTVGFGDYVA
1250	2600	A	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF EWVYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA
1251	2601	A	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	A	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	A	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGULLTFEFHPRSKLHL
1255	2605	A	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT
1256	2606	A	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE
1257	2607	A	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

Solution   Solution	SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
1258	eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
	uence		Ì	914			
Peptide   Page		ļ					
1258		1				sequence.	
							nucleotide insertion
GARLI-PGHI-PSERPPRI-IT/OQPPAAAFROPY-PQGGGHIHPI-PTGQFPCAYSGGSGALILIS	1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH
VIGECOSSYVTGAACISPUTCREWFEADT							GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR
FRSRSLREFEEALFCHTKSFPISWDAYWDRND   PLRDVDEAAVPULCICSADPOEPPDHILTT   ELFHSNPYFFLLISRHGGHCGFPROFIPTHILTT   ELFHSNPYFFLLISRHGGHCGFPROFIPTHILTT   ELFHSNPYFFLLISRHGGHCGFPROFIPTHILTT   ELFHSNPYFFLLISRHGGHCGFPROFIPTHILTT   ELFHSNPYFFLLISRHGGHCGFRREFPAWS   HEVILESFRALTEFFRTEERIKGSRIRASFI.G   GRRRGGALQRREVSSSSNI.EEINWKRSYTRL   MAAAAGAAAAPGSRPEQDPEGAGHFOFR   YYRHPERWILRPEAFLOPLRTRAPSAEDSQR   ERPAARSOPERWYRYFWAAVLAPYLALSQD   PMYKSSASOQOASGSYNHVAGAGGGGAFIVL   PLAK   SKVYWDSSFCAVNRKLAIVAGGGGAFIVL   PLAK   SKVYWDSSFCAVNRKLAIVAGGGGAFIVL   PLAK   SKVYWDSSFCAVNRKLAIVAGGGGAFIVL   PLAK   GREGICST   PRYPRORSHRIKKUSTK   CORSCICSAEKDGGDVKALYRRSQALEKIGR   LOQAVLDQRCVSLEPKNKYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLUARDA   LOQAVLDQRCVSLEPKNYFQFALRNIGGG   LOQAVLDQRCVSLUARDA   LOQA							YLGECGSSSYVTGAACISPVLRCREWFEAGLP
ELFHSNYFFILLISRHGGHCGFLRQEPLPAWS							
HEVILESTRALTEFFRIERIGLISRIRASPIG GRRRGGALOREVSSSNILEEINWKRSYTRL     MAAAAGAAAPGSREPQDRPECGAGHPGPR   YYMPERWILIPREAFLOPINFASADSOR   ERRAARSOPEMRVRYPVAAVLAPYLALSQD   PMVKSSASQQASGSYNHVREEMIAGGA   MSRRVVRQSKFRHYPQAAKADQAYEDIRV   SKYTWDSSCAVNKFLAIIVEAGGGGAFIVL   PLAK     1259   2609   A   9919   693   935   GCFKFIGESTCCWIPSSVTTQCVVAKAPRAA   TLSKAERLRSQPOPEQGGSSYRPITTAAALL   PROSCLCSAIEKBOGDWALYRRSQALEKLOR   LDQAVLDLQRCVSLEPSNVTQCDALRNIGGQ   IQEKVRYMSSTDAVEQMFQULDPEKQTE   KQKASQNLVVLAREDAGAEKIFRSNGVQLL   QRLDMGETDLMLAALRTLVGICSERDGRTV   ATISLICTRIVYSILGVESQAVSLAACHLLQV   MFDALKEGVKKOFROKEGAIIV   GREDPYRGFTLHQPKPPEGGFGGSGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGGFGGAGGG   RGKDPYRGFTLHQPKPPEGGGGGAGGAAAVPEKAKPPTEGGPGAGGG   RGKDPYRGFTLHQPKPPEGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	[		[		[		
GRRRGGALQREVSSSSNLEEIFNWKRSYTRL		}	ļ				ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS
MAAAAGAAAPAGSERPORPECGAGHPGPR		ľ			ļ		
YYRHPERWILIRPEARJEDGOR	1				Ì		
PMVKSSASGQQASGSYNIVREEMLIKAGGA   MSRRVVRQSKFRHVFGQAAKADQAYEDIRV   SKVTWDSSFCAVNFKFLAIIVFAGGGAFIVL   PLAK   SKVTWDSSFCAVNFKFLAIIVFAGGGAFIVL   PLAK   GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA   TLSKAERLRSQPGPEQGSSYRPRTPTAAAIL   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   PPRPRGRSHRKRKLVSI   GEKVRYMSSTDAKVEGMFQLLDPEKGTE   KKQKASQNLVVLAREDAGAEKIFRSNGVQLL   QREVARYMSSTDAKVEGMFQLLDPEKGTE   KKQKASQNLVVLAREDAGAEKIFRSNGVQLL   QRLDMGETDLMLAARTLVGICSEHQSRTV   ATLSILGTRRVVSILGVESQAVSLACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRAVVSILGVESQAVSLACHLLQV   MFDALKEGVKKGFRGKEGAIIV   PTRVPLHNOALLAFSPPPORQRRGTGATAES   RLFYKEASPSTHFLNLTRSSRLLAGHVSVEY   WTREGLAWGRADPHPORQRRGTGATAES   RLFYKEASPSTHFLNLTRSSRLLAGHVSVEY   WTREGLAWGRADPHFORQRGTGATAES   RLFYKEASPSTHFLNLTRSSRLLAGHVSVEY   WTREGLAWGRADPHFORQRGTGATAES   RLFYKEASPSTHFLNLTRSSRLLAGHVSVEY   WTREGLAWGRADPHFORGAGAAAVIPGLAILWAVGLGGPPPA   PPRLPFCLQELQGRHALHTFSLEGLEW   ADAGEGRLHVGAQDLATWHTLSPLGLW   ADAGEGRLHVGAQDLATWHTLSPLATMAGA   ADAGEGRLHVGAQDLATWHTLSPLATMAGA   ADAGEGRLHVGAQDLATWHTLSPLATMAGA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGAGACHA   ADAGEGRLHVGA					ļ		YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR
1259   2609   A   9919   693   935   GCFKFIGESTICKUIFPSSYTTQCVVAKAPKAL   1259   2609   A   9919   693   935   GCFKFIGESTICKUIFPSSYTTQCVVAKAPKAL   1260   2610   A   9921   455   1082   QRSCLCSAIEKDGGDVKALYRRSQALEKLGR   1260   LOQAVIDLQRCVSLEPKNKVPQEALRNIGGQ   IQEKVRYMSSTDAKVEQMFQILLDPEKGTE   KQKASQNLVVLARDAGKEIFRSNGVQLL   QRLLDMGETDLMLAALRTLVGICSEHQSRTV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKOFRGKEGAIIV   GFRGAEAPGAAQAPKKKKPRFTEGGFGAGSG   GKDPYROFTILHDPKPDEFLSSLESYEIAF   PTRVDHNGALLAFSPPPPORQRRGTGATAES   RLFYKEASPSTHFLINLTRSSRLLAGHVSVEY   WTREGLAWQRADRPHCLYA   AMBORRAGAAAVIPGLALLWAGLGGPPPA   PFRLPFCLQELQGRIALHTFSLERTCSYQDEL   WADEGRLLHVGAQDLATWHTLSFLGLW   WADEGRLLHVGAQDLATWHTLSFLGLW   A   9931   168   435   AAEMGRAGAAAVIPGLALLWAGLGGPPPA   PFRLPFCLQELQGRIALHTFSLERTCSYQDEL   WADEGRLLHVGAQDLATWHTLSFLGLW   WASATSVDQRPKQGNKVSVQNGSHQKDG   CNDDDFEPYLRSPDNQSNSYPPMSDPYMFQY   YAPSIGFPYSLGEAWSQL   1264   2614   A   9941   61   277   ESIGLTALGPRRPPWEHWSDPITLKMKGWG   WLALLLGALLGTAWARRSQDLHCGACKAVR   RRVRQFNIYDG   WADEGRLLHVGKGATHLHCTV   TSILLPUGPVILWFRGVGPGRELIYNOKEGHP   RVTTVSDLTKRNNMDFSRISSITPADVGTYY   CVKRRKGSPDHVEFKSGAGTELSVRGEYSVG   PLSQVWWNSSHPYMSDPTTLKMKGWG   PLSQVWWNSSHPMSDPTTLKMKGWG   PLSQVWWNSSHPMSDPTTLKMKGWG   PLSQVWWNSSHPMSDPTTLKMKGWG   PLSQVWWNSSHPMSDPTTLKMCKG   PLSQVWWNSSHPMSDPTTLKMCKG   PLSQVWWNSSHPMSDPTTLKMCKG   PLSQVWWNSSHPMSDPTTLKMKG   PLSQVWNSSHPMSDPTTLKMCKG   PLSQVWWNSSHPMSDPSSBGSG   PLSQVWWNSSHPMSDPSSBGSG   PLSQVWWNSSHPMSDPSSBGSG   PLSQVWWNSSHPMSDPSSBGS		}		į			
SKYTWDSSFCAVNPKFLAIIVEAGGGGAFIVL   PLAK					[		
1269   2609   A   9919   693   935   GCFKFIGESTCWIFPSSVTTQCVVAKAPRAA   TLSKAERLRSQPGEQGGSYRPRIPTIAAAIL   PPRFGRSHRKRKLVSTK   PPRFGRSHRKRKLVSTK   PPRFGRSHRKRKLVSTK   CRSCLCSAIERDGGDVKALYRRSQALEKLGR   LDQAVLDLQRCVSLEPKNKVPGEALRNIGGQ   IQEKVRYMSSTDAKVEQMFQILLDPEERGTE   KKQKASQNLVVLAREDAGAEKIFRSNOVQLL   QRLLDMGETDLMAALRTLVGICSEHQSRTV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRQEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRQFTGGFGAGSG   RGKDPYRGFTLLHQFKPPKDEFLSSLESYEIAF   PTRVDHNGALLAFPPPQRGRRGTGATAES   RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY   WTREGLAWQRADRPHCLYA   WTREGLAWQRADRPHCLYA   AAEMGRAGAAAVPROGAGPPA   PPRLPFCLQELQGRHALHTFSLERTCSYQDFL   WADEGRLLHVQAQDLATWHTLSPLGLW   WADEGRLLHVQAQDLATWHTLSPLGLW   WADEGRLHVQAQDLATWHTLSPLGT   WADEGRLHVQAQDLATWHTLSPLGT   WADEGRLHVQAQDLATWHTLSPLGT   WADEGRLHVQAQDLATWHTLSPLGT   WADEGRLHVQAQDLATWHTLSPLGT   WADEGRLHVQAQUAWARSQU   VASIGNAWARSQU   WATEGRLAWGAGAR   WAT				Į	}		
1260   2610   A   9921   455   1082   ORSCICASIEKOGOVKALYRRSQALEKLOR   CRECICASIEKOGOVKALYRRSQALEKLOR   LDQAVLDLORCVSLEPKNKVYGEALRNIGGQ   IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE   KQKASQNLVVLAREDAGAEKIFRSNOVQLL   ORLLDMGETDLMLAALRTLVGICSEHQSRTV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   WFDALKEGVKKKFRFTEGFGAGSG   RGKDPYRGFTLLNITRSSRLLAGHVSVEY   WTEGLAWQRAGPKFROTELJALAGHVSVEY   WTEGLAWQRAGPHFLUYA   ASEMGRAGAAAVIPGLALLWAVGLGGPPPA   PPRLPFCLQELQQRHALHTFSLERTCSYQDFL   WADEGRLLHVGAQQLATWHTLSPLGLW   WADEGRLLHVGAQGRATHATHOTV   TSLLPVQFVLWFRGVGYGRELLYNQKEGHFP   RVTTVYSDLTKNNNMDFSIRISTIPADVGTTYY   CVKTRKGSPDHVEFKSGAGTELSVRGEYSVG   FLSQVWWLSSHPFMN   PKNNACHLIFTAVCQPFCKHGECIGPNKCKC   PGYAGKTONQGKTV   PGYAGKTONQGKTV   PGYAGKTONQGKTV   PGYAGKTONQGKTV   PGYAGKTONQGKTV   PGYAGKTONQGKTV   PGYAGKTONQGKTV   PGPGGGFASKTKKHPQQKVKVFRAADPLV   GVFLWGVAHSINELSQVPPPVMLLPDDFKAS   SKIKVNNSHIFIRENLPSHTKKEVCPQVFRNL   RDFGDDQDVLVSTRNPPSESGSGGRLIS   YDRILVIKEVSSEDIADMHSNLSNYHQVRPLS   SPILSLSLLTYSSAIVSNRCQCJGRKLIGRENP   PGGGFGASKTKKUQCJGRKLIGRENP   PGGGFGASKTKKUGQGRCUGRKLIGRENP   PGGGFGASKTKKUGQGRCUGRKLIGRENP   PGGGFG	1250	2400		0010	602	025	
PPRPORSHRKRKI.VSTK	1239	2009	A	9919	093	. 933	
LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ   IQEKVRYMSSTDAKVEQMFQILLDFEEKGTE   KKQKASQNLVVLAREDAGAEKJFRSNGVQLL   QRLLDMGETDLMLAALRTLVGICSEHQSRTV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGVESQAVSLAACHLQV   MFDALKEGVKKKFRPTEGGFGAGSG   RCKDPYRGPTLLHQFKPFDEFLSSLESYEIAF   PTRVDINGALLAFSPPPPORQRRGTGATAES   RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY   WTREGLAWQRADRPHCLYA   WTREGLAWQRADRPHCLYA   ARMGRAGAGAAVQGLATWHTLSPLGLW   WADERILHVGAQDLATWHTLSPLGLW   WADERILHVGAQDLATWHTLSPLGLW   WADERILHVGAQDLATWHTLSPLGLW   WADERILHVGAQDLATWHTLSPLGLW   WADERILHVGAQDLATWHTLSPLGLW   YAPSIGFPYSLGEAAWSQL   SEIGLTAIGPRRRPWEHRWSDPTILKMGWG   CNDDDEFPYLRSPDNQSNSYPPMSDPYMPGY   YAPSIGFPYSLGEAAWSQL   YAPSIGFPYSLGEAWSQL   SEIGLTAIGPRRRPWEHRWSDPTILKMGWG   WLALLLGALGTAWARRSQDLHCGACKAVR   RRVRQFNIYDY   TSLLPVGPVLWFRQVGPGRELLTVNQKEGHFP   RVTTVSDLTKRNNMDFSRISISTPADVGTTYY   CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   FLSQVWWWLSSHPFMN   RDRFGIDDQDVLVSLTRNPPSESGSGGRFLIS   SPLSSSLTYSSAVSRCQLGRKLIGRENP   GVFLWGVAHSINELSQVPPPVMLLPDDFKAS   SKIKVNNHLFHRENLPSHKKFKEYCPQVFRNL   RDRFGIDDQDVLVSLTRNPPSESGSGGRFLIS   YDRTLVKEVSSEDIADMHSNLSNYHQVRPLS   SPLSSSLLTYSSAVSNRCQLGRKLIGRENP   FLSQVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPPHRAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPPHRAVLSP   FLSGVLFVPSNGVPAVCHMVGRPPHRAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP   FLSGVLFVPSNGVPAVCHMVLGRPHFAVLSP		<b> </b>			<u> </u>		PPRPGRSHRKRKLVSTK
IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE   KKQKASQNLVVLAREDAGAEKIFRSNGVQLL   QRLLDMGETDLMLAALRTLVGICSEHQSRTV   ATLSILGTRRVVSILGYESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGYESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGYESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVVSILGYESQAVSLAACHLLQV   MFDALKEGVKKGFRGKEGAIIV   ATLSILGTRRVAVSILGTRATAES   RGKDPYRGFTLLHQPKPPKDEFLSSLESYEIAF   PTRVDHNGALLAFSPPPPPQRRGTGATAES   RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY   WTREGLAWQRADRPHCLYA   WTREGLAWQRADRPHCLYA   ALEMGRAGAAVIPGLALLWAVGLGGPPPA   PPRLPFCLQELQGRHALHTFSLERTCSYQDFL   WADEGRLHVGAQDLATWHTLSPLGLW   WADEGRLHVGAQDLATWHTLSPLGLW   WADEGRLHVGAQDLATWHTLSPLGLW   CNDDDFEPYLRSPDNQSNSYPPMSDPYMFGY   YAPSIGFPYSLGEAWSQL   CNDDDFEPYLRSPDNQSNSYPPMSDPYMFGY   YAPSIGFPYSLGEAWSQL   SIGUTALGFRRPWEHRWSDPITLKMKGWG   WLALLLGALLGTAWARRSQDLHCGACKAVR   RRVRQFNIYDY   RVSEYSKMPVPASWPHPPGPFLLTILLGUT   EVAGEEELQMIQPEKLLLVTVGKTATLHCTV   TSLLPVGPVLWFRGYGPRELIYNQKEGHFP   RVTTYSDLTKRNNMDFSRISSITPADVGTYY   CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG   FLSQVWWLSSHPPIMN   FLSQVWWWLSSHPPIMN   PRNNACHLLFTAVCQPRCKHGECIGPNKCKC   HPGYAGKTCNQGRKTV   GVFLWGVAHSINELSQVPPPVMLLPDDFKAS   SKIKVNHLFHRENLPSHEKKEKYCPQVFRNL   RDRFGIDDQVI/SLTRNPPSESSEGSDGRFLIS   YQRFLUKGVSSEDIADMHSNLSNYHQVRPLS   SPILSSLTYSSAIVSNRCQLGRKLIGRENP   1268   2618   A 10005   2 209   GEGYELFVPSNGVPAVCHMVGRRPHERVLSP   1268   2618   A 10005   2 209   GEGYELFVPSNGVPAVCHMVGRRPHERVLSP   1268   2618   A 10005   2 209   GEGYELFVPSNGVPAVCHMVGRRPHERVLSP   1268   2618   A 10005   2 209   GEGYELFVPSNGVPAVCHMVGRRPHERVLSP   1268   2618   A 10005   2 209   GEGYELFVPSNGVPAVCHMVGRRPHERVLSP   1268   2618   A 10005   2 209   GEGYELFVPSNGVPAVCHMVGRRPHERVLSP   1268	1260	2610	A	9921	455	1082	
KKQKASQNLVVLAREDĀGĀĒKIFRSNGVQLL QRLLDMGĒTDLMLARITLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLQV ATLSILGTRRVVSILGVESQAVSLAACHLLQV MFDALKEGVKKGFRGKEGAIIV GFRGAEĀGGAQAPKKKKPPTĒGĢĀGAĞĀ RGKDPYRĢĪTLHQPKPPKDĒFLSSLESYĒIAF PTRVDHNGALLAFSPPPPQRQRRGTGATĀES RLFYKEASPSTHFLLNIRSSRLLAGHVSVEY WTRĒGLAWQRĀDRPHCLYĀ PRRLPTRĒGLĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀGĀ							
1261   2611   A   9928   1   438   GFRGAEAPGAAQAPKKKKPPTEGGPGAGSG RGKDPYRGPILLHQPKPPKDEFLSSLESYEIAF PTRYDHINGALLAFSPPPPQRORRGTGATAES RLFYKEASPSTHFLINLTRSSRLLAGHYSVEY WTREGLAWQRADRPHCLYA     1262   2612   A   9931   168   435   AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRAPPROGRAGA   PPRAPPROGRAGA   PPRAPPROGRAGA   PPRAPPROGRAGA   PPRAPPROGRAGA   PRAPPROPPA     1263   2613   A   9938   247   488   RMSATSVDQRFKGGNKVSVQNGSIHQKDG CNDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGPPYSLGEAAWSQL     1264   2614   A   9941   61   277   ESIGLTALGPRRRPWEHRWSDPITLKMKGWG WLALLLGALGTAWARSQDLHCGACKAVR RRVRQFNIYDY     1265   2615   A   9956   2   522   FVASEVSKMPVPASWPHPPGFFLLITLLGLT EVAGEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNNDFSRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWLSSHPPMN     1266   2616   A   10002   243   387   PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV     1267   2617   A   10004   36   707   LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINLESQVPPPVMILPDDFKAS SKIKVNNHLFHRENLPSHFKKEYCPQVFRNL RDRFGIDDQYLVSLTRNPPSESEGSOGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILLSSLLTYSSAIVSNRCQLGRKLISNYHGVRPLS SPILLSSLLTYSSAIVSNRCQLGRKLIGRENP     1268   2618   A   10005   2   209   GEGYELFYSNGVPACMINGREPIRAVLSP					•		
1261   2611   A   9928   I   438   GFRGAEAPGAAQAPKKKPPTEGGFGAGSG   RGKDPYRGPTLLHQPKPPKDEFLSSLESVEIAF   PTRVDHNGALLAFSPPPPQRQRRGTGATAES   RIFYKEASPSTHFILNLTRSSRLLAGHVSVEY   WTREGLAWQRADRPHCLYA     1262   2612   A   9931   168   435   AAEMGRAGAAAVIPGLALLWAVGLGGPPPA   PPRLPFCLQELQGRHALHTFSLERTCSYQDFL   WADEGRLLHVGAQDLATWHTLSPLGLW     1263   2613   A   9938   247   488   RMSATSVDQRFKQGGNKVSVQNGSIHQKDG   CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY   YAPSIGFFYSLGEAWSQL     1264   2614   A   9941   61   277   ESIGLTALGPRRPWEHRWSDPITLKMKGWG   WLALLGALLGTAWARRSQDLHCGACKAVR   RRVQCPNIYDY   CYKFRKGSPDHYEFKSGAGTELSVRGEYSVG   FLSQVWWWLSHPFMN     1265   2615   A   9956   2   522   FVASEVSKMPVPASWPHPPGFFLLLTLLGLT   EVAGEEELQMQPEKLLLVTVGKTATLHCTV   TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP   RVTTVSDLTKRNNMDFSIRISSTTADVGTYY   CYKFRKGSPDHYEFKSGAGTELSVRGEYSVG   FLSQVWWWLSSHPFMN     1266   2616   A   10002   243   387   PKNNACHLLFTAVCQPRCKHGECIGPNKCKC   HPGYAGKTCNQGRKTV   GVFLWGVAHSINELSQVPPPVMLLPDDFKAS   SKIKVNNHLFHRENLPSHFKFKEVPQVFRNL   RDRFGIDDQDYLVSLTRNPPSESEGSDGFFLIS   YDRTLVIKEVSSEDIADMISNLSNYHQVRPLS   SPILSSLSLTYSSAIVSNRCQLGRKLIGRENP   1268   2618   A   10005   2   209   GEGYELIFVPSNGVPAVCHMYGRRFHRAVLSP							QRLLDMGETDLMLAALRTLVGICSEHQSRTV
1261		ĺ					ATLSILGTRRVVSILGVESQAVSLAACHLLQV
RGKDPYRGPTLHOPKPPKDEFLSSLESYEIAF   PTRVDHNGALLAFSPPPPQQRRGTGATAES   RLFYKEASPSTHFILNLTRSSRLLAGHVSVEY   WTREGLAWQRADRPHCLYA	1261	2611	A	9928	1	438	
RIFYKEASPSTHFILNLTRSSRLIAGHVSVEY   WTREGLAWQRADRPHCLYA							RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF
WTREGLAWQRADRPHCLYA							PTRVDHNGALLAFSPPPPQRQRRGTGATAES
1262   2612		]	]				
PPRLPFCLQELQGRHALHTFSLERTCSYQDFL WADEGRLLHVGAQDLATWHTLSPLGLW WADEGRLLHVGAQDLATWHTLSPLGLW RMSATSVDQRPKQQGNKVSVQNGSIHQKDG CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL  1264 2614 A 9941 61 277 ESIGLTALGPRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVRQFNIYDY  1265 2615 A 9956 2 522 FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT EVAGEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP 1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGGRFFHRAVLSP	1262	2612	A	9931	168	435	
1263	l		1				PPRLPFCLQELQGRHALHTFSLERTCSYQDFL
CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL  1264 2614 A 9941 61 277 ESIGLTALGPRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVRQFNIYDY  1265 2615 A 9956 2 522 FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESGGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1263	2613	Δ	0038	247	488	
YAPSIGFPYSLGEAAWSQL     1264   2614   A   9941   61   277   ESIGLTALGPRRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVRQFNIYDY     1265   2615   A   9956   2   522   FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWLSSHPFMN     1266   2616   A   10002   243   387   PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV     1267   2617   A   10004   36   707   LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP     1268   2618   A   10005   2   209   GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1203	2013	^	7730	447	700	
WLALLEGALEGTAWARRSQDLHCGACKAVR RRVRQFNIYDY  1265 2615 A 9956 2 522 FVASEVSKMPVPASWPHPPGPFLLITLLEGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP							YAPSIGFPYSLGEAAWSQL
RRVRQFNIYDY  1265 2615 A 9956 2 522 FVASEVSKMPVPASWPHPPGPFLLTLLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1264	2614	A	9941	61	277	
1265 2615 A 9956 2 522 FVASEVSKMPVPASWPHPPGFFLLTLLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP							
TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1265	2615	A	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT
RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP							
CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN  1266 2616 A 10002 243 387 PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP			}				
FLSQVWWWLSSHPFMN     1266   2616	1						CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG
HPGYAGKTCNQGRKTV  1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKE YCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP  1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1000	0.00	<u> </u>	10055	0.10		FLSQVWWWLSSHPFMN
1267 2617 A 10004 36 707 LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP 1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1206	2016	<b>A</b>	10002	243	387	
PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP 1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG
SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP 1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP							PGPGFGFASKTKKKHFVOOKVKVFRAADPLV
RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP 1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP	1						SKIKVNNHI FHRENI PSHEKEKEVCDOVERNI
YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP 1268 2618 A 10005 2 209 GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP			1				RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS
1268   2618   A   10005   2   209   GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP							YDRTLVIKEVSSEDIADMHSNLSNYHOVRPLS
SODELEHSLGESAAOGAAGVVI WVSWFNTD	1268	2619	Δ	10005	2	200	SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
	1200	2010	\ ^	10002	_	209	SQDELEHSLGESAAQGAAGVVLWVSWENTR

	1 000 000	157	1 650	D 11. 1		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ļ	1	Í		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ			ļ	peptide		/=possible nucleotide deletion, \=possible
	l		1	sequence		nucleotide insertion
			<u> </u>			TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD
ł		ł	l	Ì		LQLRNLSVADHSKTQVQKKENKSLKRDTKAI
		1	1	1		IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL
	Į.	1	İ		ļ	AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
)	ļ	1	ļ		}	VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
		1				PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII
		1			}	SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR
1	ŀ					RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK
ŀ	i	l	ł	1	ł	QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
		<b>!</b>			}	QPKLDRTSSFRQILPRFRSADHDRYRGWSMW
1		1	1.	1		DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP
12.7	702.	1	100.5		1 503	FLSGAEVSQSCRRRGGRA
1272	2622	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT
12/2	2022	] ^	10014	<b>\</b> '	1 300	LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR
Į.	1	İ				LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT
1	j	1			] _	SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	1339	
12/3	2023	A	10010	1 1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV
	1		1	Ì		SAPRRAASGPSGSAPAVAAAAAQPGSYPALS
1	1	1				AQAAREPAAFWGPLARDTLVWDTPYHTVW
1	1				1	DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
1	1		l	į.		PESVALIWERDEPGTEVRITYRELLETTCRLA
l	1		1	<b>'</b>	1	NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
1	1		1			CARIGAVHTVIFAGFSAESLAGRINDAKCKVV
		1				ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH
1		1	1		l	VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
			1			VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT
ľ	1		1			QAGYLLYAALTHKLVFDHQPGDIFGCVADIG
1	ļ		i i			WITGHSYVVYGPLCNGATSVLFESTPVYPNA
					]	GRYWETVERLKINQFYGAPTAVRLLLKYGD
i .	1	1	1 .	1		AWVKKYDRSSLRTLGSVGEPINCEAWEWLH
						RVVGDSRCTLVDTWWQT
1274	2624	A	10017	1	3750	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG
İ	1		i	1		KTLGSFFGSLPGFSSARNLVANAHSSARARPA
						ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
	1			]	_	KELQPSEKMVSGAKDLVCSKMSRAKDAVSS
1 .		1	}	Į.	1	GVASVVDVAKGVVQGGLDTTRSALTGTKEV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG
1	1	1				TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV
}	] ,		1	1	1	LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
l	1		1			KAVLTGTKDAVSTGLTGAVNVARGSIOTGV
1	}			}	ļ	DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT
1					Į.	GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT
				1	1	IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA
J	}		ŀ	1	1	KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN
		1	1			LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG
	1			ļ	1	AANVAKGAMQTGLNTTQNIATGTKDTVCSG
1				1		VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC
1	1	1		ĺ	-	
		1		1		SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
1		1		1.	ł	AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG
}	}	1	}	1	1	TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV
1			]	1	1	VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
		1		1		KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM
		1		1		DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT
1	1		1	1		GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
1	1	l	1	1	}	QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK
1		1		1		GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN
				<del></del>		<del>*</del>

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion  VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA
						VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV TGAANVAKGAVQMGVDTAKTVLTGTKDTV CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT GTKDAVSTGLTGAVNLAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG VDASKAVLMGTKDTVFSGVTGAMSMAKGA VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA TGATHTGLSTFQNWLPSTPATSWGGLTSSRT TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP AWEAAATTKGLATDVATFTQGAAPGREDTG LLATTHGPEEAPRLAMLQNELEGLGDIFHPM NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL QDCFRL
1275	2625	A	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM YKRKYSAANTKVEKKKKEKVLAPVTKPVGG DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK CRIPNNSLPYFHKRPQARMLLLALFCVAVSV VWGVFRNEDQ
1277	2627	A	10035	51	869	YSRFTVPLPATMASSEVARHLLFQSHMATKT TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT LFNTNFEDYESSHFCPNVKLKAQTYELQESN VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE SSGQSAQQPYLPINSPPHRLADVADVHLFSSV LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST VMGAVGESLSVQCRYEEKYKTFNKYWCRQP CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL TFTVTLENLTADDAGKYRCGIATILQEDGLSG FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS FLRGLFGGNTRIEEACEMYTRAANMFKMAK NWSAAGNAFCQAAKLHMQLQSKHDSATSFV DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF YEFQLTAVSEGGVLSESSTANITVVASDSPY GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR LWYKTMSGTAEAGLDFVPAAGELLFEAGEM RKSLHVEILDDDYPEGPEEFSLTITKVELQGR GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN
1283	2633	A	10088	316	516	SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT
.203		"	10000			VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634		10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	A	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	A	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

SEQ ID NO: of nucleotide peptide confide sequence peptide sequence unice of the sequence peptide sequence of the sequence peptide sequence pep							
mucleoide seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence	nucl-	nentide	<u> </u>	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
19496							
1914   ng no first smino acid residue of peptide pep				1			
mimo acid residue of peptide residue of peptide residue of peptide sequence   T-T-Tarronine, V-Valiae, W-Typtophan, Y-Tyrosine, N-Valnow, N-Supo codon, P-possible nucleotide deletion, V-possible nucleotide deletion, V-possible nucleotide insertion   LGNVLTSTPNAKTVNGKASSSDSGASSEGEE		uence					
residue of   popular   p	uence		i	914			
Peptide   Possible nucleotide deletion,   -possible nucleotide insertion	ŀ	Ì	ì	į	amino acid	of peptide	
Peptide   Possible nucleotide deletion,   possible   paquence		ŀ		ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	1	1	ļ	]			
1293			1				
1293		<u> </u>		<u> </u>	sequence		
193	l -	1		1		ŀ	
	1		]	Į.			AC
	1293	2643	Δ	10124	2	989	PLMSLVRVVEEVAASSAOKTPSRLENVYMVC
CGRGIRDS ARMCSTCACVÉTYGKALEKULK   GVKINCHIGKMKYRKNIKIMERIPERIALOSOILV   CTDVMARGIDIPEVNIVULQYDPPSNASAFUL   CTDVMARGIDIPEVNIVULQYDPPSNASAFUL   CTDVMARGIDIPEVNIVULQYDPPSNASAFUL   RCGRTARIGHGGSAL VFLLPMERSTNIFLAIN   QKCPLQEMKPQRNTADLLFKLK SMALADRA   VFEKGMKAPVSYVQAYAKHECNLFKLDL   DFASLARGFALLRMFKMPELREKGRGFDFVPV   DVNTIDTIPFKDKIREKGRGKGLEQGREKKTEN   EGRRIFKIKNASWKQKAKKK   CKREMIT   CHARLES     LIGHVYTILLLAFLAFLINGCOSQWINVLETQ   LLFLLSVLGIFGLAFAFIELNQGTAPVRYFL   GVLFALCPSCLLAHASNLVKLVRGCVSFSWT   TILCIAIGCSLQQIMATEYVTLIMTRGMMFVN    MTPCQLNVDFYVLLVVVLEHALFVSKAT   FCGPCENWKQHGRLIFITVLFSIIWVVWSIKL   LKRWPGFQRQPWDDPVCLALVTNAWVFL   LLYIVPELCILLYRSCRGECPLQGNACPVTAYQ   HSQVPNOELSRDKWVLLINSDPLSHSGA   LIGHVYTHNSQWCFLFQDIPFGWLFQSGAPG   GRGAPRQEGFGSSWRQV   1295	1 12/3	2043	'A	10127	-		
GVKINCHIGKMKYKRNIKIPERRILOSGIU	i	1		1	l		
CIDVMARGIDIPEVNIWUQYDPPSNASAVI   RCGRTARIGHGGSALVFILMERIKSMALAND   QKCPLQEMKPQRNTADLIPKLKSMALAND   QKCPLQEMKPQRNTADLIPKLKSMALAND   VFEKGMKAPSVYQAVAKHERICHIFRIKDL   DFASLARGFALLRMPEMPELRGKOFPDFVPV   DVNTDTIPFKDKIREKORQKILEQQRREKTEN   EGRREFIKNRAWSKQKAKKK   EGRREFIKNRAWSKQKAKKK   LIGVVYILLILAFFLMRKIQDCSWNVLPTQ   LLFLLSVLGIFGLAFAFIELNQQTAPVRYFIC   GVLFALCPSCLLAHASNLVKLVRGCVSTSWT   TILCIAIGCSILQIIIATEYVTLIMTRGMMFVN   MTPCQLNVDFVVLLVYVLFALKRGCVSTSWT   TILCIAIGCSILQIIIATEYVTLIMTRGMMFVN   MTPCQLNVDFVVLLVYVLFALKRGCVSTSWT   TILCIAIGCSILQIIIATEYVTLIMTRGMMFVN   HSQVPNQELSRDKWKVLINSDFLSHSQA   HSQVPNQELSRDKWKVLINSDFLSHS   HSQVPNQELSRDKWKVLINSDFLSDFLSDFLSHS   HSQVPNQELSRDKWKVLINSDFLSDFLSDFLST   HSQVPNQELSRDKWKVLINSDFLSDFLSDFLST   HSQVPNQELSRDKWKVLINSDFLSDFLSDFLST   HSQVPNQELSRDKWKVLINSDFLSDFLSDFLSDFLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   HSQVPNQELSRDKASAMSSPLST   H	1	1	1	1	Į.		
RCGRTARIGHGGSALVELMEESTINTALIN   QKCPLQEMKQRNTADLIPKLKSMALADRA   VFEKGMKAFVSYVQAYAKHECHLIFELKDL   DFASLARGFALLRMFKMPELRGKQFPDFVPV   DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN   EGRRKIRNKAWSKQKAKKK   EGRRKIRNKAWSKQKAKKK   ULGIVVIILLLAFLFLMKRIQDCSQWNJLTQ   LIPLISVLGIFGLAFAFIELINQTAPVRYELF   GVLFALCFSCLLAHASNLVKLVRGCVSSSWT   TILCIAIGCSLQIAITAFVYLINGMMFVN   MTPCQLNVDFVVLLVYVLFLMALTFVSKAT   FCCPCENWKOMGRLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCCWSCCCPC   FCCPCCCWSCCCPC   FCCPCCCWSCCCPC   FCCPCCCPTCCRTSCCCTTCCTRYSCCCWSCCCPC   FCCPCCCWSCCCPTCCRTSCCTTCCTRYSCCCWSCCCPC   FCCPCCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCCWSCCCPTCCTTCCTTCTCTTCTCTTCCTTCTCCTTCTCTTCT		i			1		GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV
RCGRTARIGHGGSALVELMEESTINTALIN   QKCPLQEMKQRNTADLIPKLKSMALADRA   VFEKGMKAFVSYVQAYAKHECHLIFELKDL   DFASLARGFALLRMFKMPELRGKQFPDFVPV   DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN   EGRRKIRNKAWSKQKAKKK   EGRRKIRNKAWSKQKAKKK   ULGIVVIILLLAFLFLMKRIQDCSQWNJLTQ   LIPLISVLGIFGLAFAFIELINQTAPVRYELF   GVLFALCFSCLLAHASNLVKLVRGCVSSSWT   TILCIAIGCSLQIAITAFVYLINGMMFVN   MTPCQLNVDFVVLLVYVLFLMALTFVSKAT   FCCPCENWKOMGRLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCENWKOMFLIFTUTSKAT   FCCPCCWSCCCPC   FCCPCCCWSCCCPC   FCCPCCCWSCCCPC   FCCPCCCPTCCRTSCCCTTCCTRYSCCCWSCCCPC   FCCPCCCWSCCCPTCCRTSCCTTCCTRYSCCCWSCCCPC   FCCPCCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCWSCCCPTCCRTSCCTTCCTRYSCC   FCCPCCCWSCCCPTCCTTCCTTCTCTTCTCTTCCTTCTCCTTCTCTTCT	Į	ļ	l		.]	1	CTDVMARGIDIPEVNWVLOYDPPSNASAFVH
		i		ł	1		
1294			ł				
DFASILARGFALLRMÉKMPELRGKQFPDFVPV	1	}		1	i		1 ' '
DVNTDTPFKDKIREKQRQKILEGQRREKTEN   EGRKFIFIKNA WSKQKAKKK     1294	l	}	ļ	1	ļ	ł	VFEKGMKAFVSYVQAYAKHECNLIFRLKDL
EGRRKFIKNKAWSKOKAKKK     1294   2644   A   10129   91   1042   VTMYKDCIESTGDYFILCDAEGPWGHILESLA     LIGIVYTILLLAFIFLMRKIQDCSQWNVLPTQ     LLPILLSVLGLFGLAFAFHELNQQTAPWTFLF     GVLFALGFSCLLAHASIN_VKLVRGCVSSSWT     TILCIAIGCSLLQHIATEYVTLIMTRGMMFVN     MTPCQLNVDFVVLLVYLMALTFPVSKAT     FCGPCENWKQHGRLIFITVLFSIHWVVWSML     LRGNPQFQRQWDDPVVCLALVTNAWVFL     LLYIVPELCILYRSCRGECPLQGNACPVTAYQ     HSRQVENQELSRDKWKVLLINSDFLSHSGA     GRGAPRQBGYGSSWRQV     1295   2645   A   10133   376   518   RPRVYTHINSQWCFLPQDHFDWLPGGSGAPG     GRGAPRQBGYGSSWRQV     RRVVITHSWGWCFLPQDHFDWLPGGSGAPG     GRGAPRQBGYGSSWRQV     GRGAPRQBGYGSSWRQV     WSLDPFMGIMSGQVGDLSPSQEKSLAQFRE     NIQDVLSALPNPDDYFLRWLQARSFDLQKS     EDMLRKHMEFRKQQDLANILLAWQPPEVVRL     YNANGICGHEGGSQPWHYGFBLQB     VANNGCGGHGGGGGDLCQETCCRPSCCE     TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT     CSRPSCQTTCCRTTCCRPSCCVSSCCRPQCC     1298   2648   A   10138   48   407   MVSSCGSVCSDQGGGDDLCQETCCRPSCCE     TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT     CSRPSCQTTCCRTTCCRPSCCVSSCCRPQCC     1298   2649   A   10161   393   GRSRSLABMFSQVPRTPASGCYYLNSMTPEG     GEMYLRFDOTTRRSPYRMSRLARRIQLVTKI     QCELAKEACDWLRAAGRANSQVPTLASGOPT     1299   2649   A   10161   393   PRFSELVDGRGRSVSARFGGSPSKAATVRSQPT     1299   2649   A   10161   1   393   PRFSELVDGRGRSVSARFGGSPSKAATVRSQPT     1299   2649   A   10161   1   393   PRFSELVDGRGRSVSARFGGSPSKAATVRSQPT     1300   2650   A   10162   98   391   AKIASLERIMPANYTCTRPDGDNTDFRYFIYA     VTYTGLGPGLIGNILLALWYYGYMKETKRA     VITMINLAHADLLQVLSLPLRFYYLKHDWPF     VPV   VTYTGLGPGLIGNILLALWYYGYMKETKRA     VITMINLAHADLLQVLSLPLRFYYLKHDWPF     VPV   VTYTGLGPGLGRGRGRASSISIOLQCKAGGT     RWLGKPDFGRRKRRGSPQEEGGLVSAAAR     LLCSGANRKVLVLXQNSTRPTTQQPAVHPSTP     PSRPLPQAGRCLVAPLRPPIPDWAAKTLAKSEPKR     RRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSK	1		1	1		ļ	DFASLARGFALLRMPKMPELRGKQFPDFVPV
EGRRKFIKNKAWSKOKAKKK     1294   2644   A   10129   91   1042   VTMYKDCIESTGDYFILCDAEGPWGHILESLA     LIGIVYTILLLAFIFLMRKIQDCSQWNVLPTQ     LLPILLSVLGLFGLAFAFHELNQQTAPWTFLF     GVLFALGFSCLLAHASIN_VKLVRGCVSSSWT     TILCIAIGCSLLQHIATEYVTLIMTRGMMFVN     MTPCQLNVDFVVLLVYLMALTFPVSKAT     FCGPCENWKQHGRLIFITVLFSIHWVVWSML     LRGNPQFQRQWDDPVVCLALVTNAWVFL     LLYIVPELCILYRSCRGECPLQGNACPVTAYQ     HSRQVENQELSRDKWKVLLINSDFLSHSGA     GRGAPRQBGYGSSWRQV     1295   2645   A   10133   376   518   RPRVYTHINSQWCFLPQDHFDWLPGGSGAPG     GRGAPRQBGYGSSWRQV     RRVVITHSWGWCFLPQDHFDWLPGGSGAPG     GRGAPRQBGYGSSWRQV     GRGAPRQBGYGSSWRQV     WSLDPFMGIMSGQVGDLSPSQEKSLAQFRE     NIQDVLSALPNPDDYFLRWLQARSFDLQKS     EDMLRKHMEFRKQQDLANILLAWQPPEVVRL     YNANGICGHEGGSQPWHYGFBLQB     VANNGCGGHGGGGGDLCQETCCRPSCCE     TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT     CSRPSCQTTCCRTTCCRPSCCVSSCCRPQCC     1298   2648   A   10138   48   407   MVSSCGSVCSDQGGGDDLCQETCCRPSCCE     TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT     CSRPSCQTTCCRTTCCRPSCCVSSCCRPQCC     1298   2649   A   10161   393   GRSRSLABMFSQVPRTPASGCYYLNSMTPEG     GEMYLRFDOTTRRSPYRMSRLARRIQLVTKI     QCELAKEACDWLRAAGRANSQVPTLASGOPT     1299   2649   A   10161   393   PRFSELVDGRGRSVSARFGGSPSKAATVRSQPT     1299   2649   A   10161   1   393   PRFSELVDGRGRSVSARFGGSPSKAATVRSQPT     1299   2649   A   10161   1   393   PRFSELVDGRGRSVSARFGGSPSKAATVRSQPT     1300   2650   A   10162   98   391   AKIASLERIMPANYTCTRPDGDNTDFRYFIYA     VTYTGLGPGLIGNILLALWYYGYMKETKRA     VITMINLAHADLLQVLSLPLRFYYLKHDWPF     VPV   VTYTGLGPGLIGNILLALWYYGYMKETKRA     VITMINLAHADLLQVLSLPLRFYYLKHDWPF     VPV   VTYTGLGPGLGRGRGRASSISIOLQCKAGGT     RWLGKPDFGRRKRRGSPQEEGGLVSAAAR     LLCSGANRKVLVLXQNSTRPTTQQPAVHPSTP     PSRPLPQAGRCLVAPLRPPIPDWAAKTLAKSEPKR     RRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSKEREHKRRTSTPV     ARRLSSELHVYDENKNSSK			1		ŀ		DVNTDTIPFKDKIREKOROKLLEOORREKTEN
10129	1				1		
ILGIVYTILLILAFLEMRKIQDCSQWNVLYFIL   CFILE	1004	0544		10100		1040	
LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF   GVLFALCFSCLLAHASNLVKLVRGCVSFSWT   TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN   MTPCQLNVDFVVLLYVLFLMALTFFVSKAT   FCGPCENWKQHGRLIVTSIIWVVWISML   LRGNPQFORQPWDDPVVCLALVTNAWVFL   LLYIVPELCLLYRSCGPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLLNSDFLSHSGA   1295   2645   A 10133   376   518   RPRVVTHNSQWCFLPQDIPGWLFQQSGAPG   GRGAFRQEGFGSSWCFUFQQSGAPG   GRGAFRQEGFGSSWCFUFQQSGAPG   GRGAFRQEGFGSSWCFUFQQSGAPG   GRGAFRQEGFGSSWCFUFQQSGAPG   GRGAFRQEGFGSSWCFUFQQSGAPG   SWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE   NIQDVLSALPNPDDYFLLRWLQARSFDLQKS   EDMLRKHMEFRKQQDLANILAWQPFEVVRL   YNANGICCHIDGEGSPWYHIVOSQDPKGLLL   SASKQELLRSFSRSCELLRSEGLQSKLJGKR   VEKILAFFGLEGLGJRDLWKPGIELLQE   SCHLAFFGSCELLRSFSRSCELLRSECLSKUJGK   VEKILAFFGLEGLGJRDLWKPGIELLQE   CSRSCCQTTCCRTTCCRPSCCVSSCCRPQCC   QFVCCQPTCCRPSCCETTCCHPXCC   QFVCCQPTCCRPSCCETTCCHPXCC   QFVCCQPTCCRPSCCETTCCHPXCC   QEMYLRFQQTTRRSPYRASGCYYLNSMTPEG   QEMYLRFQQTTRRSPYRASGLARHQLVTKI   QQEIEAKEACDWLRAAGFFQYAQLYEDSQFP   NIVAVKNDHDFLEKDLGEFLCRINT   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADASCDCGPGF   KGRCCELACIKVSRCTRLFSETKAFPVWEGG   VCHHV   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADASCDCGPGF   KGRCCELACIKVSRCTRLFSETKAFPVWEGG   VCHHV   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADASCDCGPGF   KGRCCELACIKVSRCTRLFSETKAFPVWEGG   VCHHV   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADASCDCGPGF   KGRCCELACIKVSRCTRLFSETKAFPVWEGG   VCHHV   ASAQLENMEGAPGRSSSISDLQGKGFEKGTG   EKHVPGVGSARHSPQASAGGSPWQRGKAQT   RWLGKPDPGKRRRRGSPSEEGLRVSAAAA   VIFMNLAIADLQVLSLPLRIPDWVAAKTLAKA   LRAPGKPWILAAPSIDLQGAFGLGPGFTAP   RTLSVEEPGVECNQLCJADAVTDPVLCLGQK   DPGVEGKHCEKEKISSKEKEHVHAKSEPSKP   ARLSESLHVVDENKNESKEREHKRTSTFV   ARLSESLHVVDENKNESKEREHKRRTSTFV   ARLSESLHVVDENKNESKEREHKRRTSTFV   ARLSESLHVVDENKNESKEREHKRRTSTFV   ARLSESLHVVDENKNESKEREHKRRTSTFV   ARLSESLHVVDENKNESKEREHKRRTSTFV   ARLSESLHVVDENKNESKEREHKRTSTFV   ARLSESLHVVDENKNESKEREHKRTSTFV   ARLSESLHVVDENKNESKEREHKRTSTFV   ARLSESLHVVDENKNESKEREHKRTSTFV   ARLSESLHVVDENKNESKEREHKRTSTFV   ARLSESLHVDENKNESKEREHKRTST	1294	2044	A.	10129	91	1042	
GULFALCESCLLAHASNI, VKLVRGCVSFSWT TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN MTPCQLNVDFVVLLYVLFILMALTFYSKAT FCGPCENWKQHGRLIFITVLFSKAT FCGPCENWKQHGRLIFITVLFSKAT FCGPCENWKQHGRLIFITVLFSIWVVWISML LRGNPGFQRQPQWDDPVCCIALVTNAWVFL LLYIVPELCIL VRSCRQECPLQGMACPVTAYQ HSPQVENQELSRDKWKVLLNSDFLSHSGA 1295 2645 A 10133 376 518 RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGPGSSWRQV	}	1	ł	i	ł	ł	
TILCIAIGCSILQIIIATEYYTILMIRGMMFVN   MTPCQLNVDFVVLLVYVILMALTFFVSKAT   FCGPCENWKQHGRLIFTVJFSIIWVWISML   LRGNPQFQRQPWDDPVCICLAYTNAWPTI-   LLYIVPELCILYRSCRECPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLINSDFLSHSGA     1295   2645   A   10133   376   518   RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG   GRGAPRQEGPGSSWRQV   GRGAPRQEGPGSSWRQV     1296   2646   A   10135   3   551   EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE   NIQDVL3ALPNPDDFLIRWLQARSFDLQKS   EDMLRKHMEFRKQQDLANILAWQPPEVVRL   YNANGICCHDGEGSPVWFIVGSQDPKGLL   SASKQELLRDSFRSCELLLRECELQSQKLGKR   VEKIIAIFGLEGLGRDLWKPGIELLQE   SASKQELLRDSFRSCELLRECELQSQKLGKR   VEKIIAIFGLEGLGLGLDLWKPGIELLQE   TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT   CSRPSCCQTTCCRTTCTRPSCCVSSCCRPQCC   GSPSCCQTTCCRTTCTRPSCCVSSCCRPQCC   GSPSCCQTTCCRTTCTRPSCCVSSCCRPQCC   1298   2648   A   10156   94   453   GGNRKSAEMFSQVPTRASGCYYLNSMTDED   GEMYLRAGFPQTTRSPSFRSMARARQLVTKI   QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP   NIVAVKNDHDFLEKDLGEFLCRRINT   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADABSCDCGPGF   KGRCCELACIKVSRPCTRLFSETKAFPVWEGG   VCHHV   ASASQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADABSCDCGPGF   KGRCCELACIKVSRPCTRLFSETKAFPVWEGG   VCHHV   ASAGLERIMPANYTCTRPDGDNTDFRYFIYA   VTYTGILGPGLIGNILALWYFYGYMKETKRA   VIFMINLAIADLQVLSLPLRIFYSTKAFPVWEGG   VCHHV   ASAGLERIMPANYTCTRPDGDNTDFRYFIYA   VTYTGILGPGLIGNILALWYFYGYMKETKRA   VIFMINLAIADLQVLSLPLRIFYSTKAFPVWEGG   TWORT   TWO		1	1		1	ł	LLFLLSVLGLFGLAFAFIJELNQQTAPVRYFLF
TILCIAIGCSILQIIIATEYYTILMIRGMMFVN   MTPCQLNVDFVVLLVYVILMALTFFVSKAT   FCGPCENWKQHGRLIFTVJFSIIWVWISML   LRGNPQFQRQPWDDPVCICLAYTNAWPTI-   LLYIVPELCILYRSCRECPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLINSDFLSHSGA     1295   2645   A   10133   376   518   RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG   GRGAPRQEGPGSSWRQV   GRGAPRQEGPGSSWRQV     1296   2646   A   10135   3   551   EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE   NIQDVL3ALPNPDDFLIRWLQARSFDLQKS   EDMLRKHMEFRKQQDLANILAWQPPEVVRL   YNANGICCHDGEGSPVWFIVGSQDPKGLL   SASKQELLRDSFRSCELLLRECELQSQKLGKR   VEKIIAIFGLEGLGRDLWKPGIELLQE   SASKQELLRDSFRSCELLRECELQSQKLGKR   VEKIIAIFGLEGLGLGLDLWKPGIELLQE   TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT   CSRPSCCQTTCCRTTCTRPSCCVSSCCRPQCC   GSPSCCQTTCCRTTCTRPSCCVSSCCRPQCC   GSPSCCQTTCCRTTCTRPSCCVSSCCRPQCC   1298   2648   A   10156   94   453   GGNRKSAEMFSQVPTRASGCYYLNSMTDED   GEMYLRAGFPQTTRSPSFRSMARARQLVTKI   QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP   NIVAVKNDHDFLEKDLGEFLCRRINT   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADABSCDCGPGF   KGRCCELACIKVSRPCTRLFSETKAFPVWEGG   VCHHV   ASASQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADABSCDCGPGF   KGRCCELACIKVSRPCTRLFSETKAFPVWEGG   VCHHV   ASAGLERIMPANYTCTRPDGDNTDFRYFIYA   VTYTGILGPGLIGNILALWYFYGYMKETKRA   VIFMINLAIADLQVLSLPLRIFYSTKAFPVWEGG   VCHHV   ASAGLERIMPANYTCTRPDGDNTDFRYFIYA   VTYTGILGPGLIGNILALWYFYGYMKETKRA   VIFMINLAIADLQVLSLPLRIFYSTKAFPVWEGG   TWORT   TWO	ļ	1		1			GVLFALCFSCLLAHASNLVKLVRGCVSFSWT
MTPCQLNVDFVVLLVYVLEMALTFFVSKAT   FCGPCENWKQHGRLIFTVLFSIIIWVVWISML   LRGNPQFQRQPQWDDFVVCIALVTNAWVFL   LLYIVPELCILYRSCRGEFLQGNACFVTAYQ   HSFQVENQELSRJKWKVLINSDFJSHGA     1295   2645   A   10133   376   518   RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG   GRGAPRQEGFGSSWRQV     1296   2646   A   10135   3   551   EWSLDPFMGMSGGVGDLSPSQEKSLAQFRE   NIQDVLSALPNPDDYFLLRWLQARSFDLQKS   EDMLRKHMEFRKQQDLANILAWQPPEVVRL   YAANGICGHDGEGSPVWTHIVGSQDPKGLLL   SASKQELLRDSFRSCELLLRECELQSQKLGKR   VEKILAFGLEGLADLWKPGIELLQE     1297   2647   A   10138   48   407   MVSSCGGSVCSDQGCGQDLCQETCCRPSCC   CSPSCCQTTCCRTTCYRSCCVSSCCRPQCC   QVCCQPTCCRPSCCSTCCCPTCCCCTTCYRSCC   QCCQVCCQPTCCRPSCCTTTCCRTTCCRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCCRTCCRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCRSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCCRPSCCTTTCYRRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCC   QVCCQPTCRSCTTCYRSCCVSSCCRPQCCQVCQPTCRSCTTCYRSCCVSSCCRPQCCQSCCVQTCCPTCCPSCCTTCCRSCCTTTCCRTTCCRTCCTTTCTTTCTTTCTTT	1	1	1			ł	
	l	1 .	<b>}</b> ,	ł	i	ł	
LRGNPQFQRQPQWDDPVVCIALVTNAWVFL   LLYIVPELCILYRSCRQECPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLINSDFLSHSGA   10133   376   518   RPRVVTINSQWCFLPQDIPGWLPGQSGAPG   GRGAPRQEGFGSWRQV   1296   2646   A   10135   3   551   EWSLDFFMGIMSGQVGDLSPSQEKSLAQFRE   NIQDVLSALPNPDDYFLLRWLQARSFDLQKS   EDMLRKIMEFRKQQDLANILAWQPEVVRL   YANANGIGGHDGEGSPVWYHIVGSQDPKGLLL   SASKQELLRDSFRSCELLRECELQSQKLGKR   VEKIIAFGLEGLGLRDLWKPGIBLQE   1297   2647   A   10138   48   407   MVSSCCGSVCSDQGGGDDLCQETCCRPSCCE   TTCRTTCCRPSCCSTCCPCCC   QPVCCQPTCCRTTCYRPSCCVSSCCRQCC   QPVCCQPTCCRTTCCPRSCCSTCCPCCC   QPVCCQPTCCRTTCCPRSCCYSCCRQCCQVCCQPT   GRPSCCQTTCCRTTCYRPSCCVSSCCRQCC   QPVCCQPTCCRTSCCSTCCPRXCC   QEMYLRPDQTTRRSPYRMSRLARHQLVTKI   QCEIFAKEACDWLRAAGFPQYAQLVEDSQFP   NIVAVKNDHDFLEKDLGEFLCRRLNT   1299   2649   A   10161   1   393   PRFSELVDGRGRVSARFGGSPSKAATVRSQPT   ASAQLEMMEBAPKRVSLALQLPEHGSKDIGN   VFONCSENPCQNGGTCVPGADAHSCDCCPGF   KGRCELACKVSRPCTRFSETKAFPVWEGG   VCHIV   VTYTGLGPGLIGNLALWYFGYMKETKRA   VTYTGLGPGLIGNCALWYFGYMKETKRA   VTYTGLGPGLIGNCALWYFGYMKETKRA   VTYTGLGPGLIGNCALWYFGYMKETKRA   VTYTGLGPGLIGNCALWYFGYMKETKRA   VTYTGLGPGLIGNCALWYFGYMKETKRA   VTYTGLGPGLIGNCALWYFGYMKETKRA   VTYTGLGPGLIGNGAGFERGTGCASSSISDLQCGGFGEGTG   EKTYPGOYGARRISPQASAGGSPWQGKAQT   RWLGKPD	ì	1 .		1			,
LLXIVPELCILYRSCRQECPLQGNACPVTAYQ			1	1	1		
HSFQVENQELSRDKWKVLLNSDFLSHSGA	{	1	1		}	Į.	LRGNPQFQRQPQWDDPVVCIALVTNAWVFL
HSFQVENQELSRDKWKVLLNSDFLSHSGA	1		1	1	1	í	LLYIVPELCILYRSCROECPLOGNACPVTAYO
1295		1		1 •		1	
1296	1005	10645	<del>                                     </del>	+10122	200	610	
1296	1293	2045	A	10133	3/0	219	1
NIQDVLSALPNPDDYFLLRWLQARSFDLQKS   EDMLRKHMEFRKQDLANILAWQPPEVVRL   YNANGICGHDGEGSPVWHIVGSQDPKGLLL   SASKQELLRDSFRSCELLLRECELQSQKLGKR   YEKIIAFGLEGLGLRDLWKPGIELLQE   MYSSCCSDQGCQDLQQETCCRPSCCE   TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT   CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC   QPVCCQPTCCRPSCCETTCCHPXCC     1298   2648	l		<u>L</u>		i		
EDMLRKHMEFRKQQDLANILAWQPPEVVRL   YNANGICGHDGESPYWYHIVGSQDPKGLLL   SASKQELLRDSFRSCELLIRECELQSQKLGKR   VEKHAIFGLEGLGLRDLWKPGIELLQE	1296	2646	A	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE
EDMLRKHMEFRKQQDLANILAWQPPEVVRL   YNANGICGHDGESPYWYHIVGSQDPKGLLL   SASKQELLRDSFRSCELLIRECELQSQKLGKR   VEKHAIFGLEGLGLRDLWKPGIELLQE	1	1	1		j	ĺ	NIODVLSALPNPDDYFLLRWLOARSFDLOKS
1297   2647   A   10138   48   407   MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE	ļ		1	1	i	1	
SASKQELLRDSFRSCELLLRECELQSOKLGKR	1 .	1		ł		1	
VEKITÄIFGLEGLGLRDLWKPGIELLQE	ļ		i	ļ			
1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCVTSCCTPQCCQSVCCQPT CSRPSCCVTTCCRTSCCVSSCCRQCC QPVCCQPTCCRTSCCVSSCCRQCC QPVCCQPTCCRTSCCTTCCHPXCC  1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRLARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSGFP INIVAVKNDHDFLEKDLGEFLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIPMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	Ι,	i	İ		l	1	SASKQELLRDSFRSCELLLRECELQSQKLGKR
TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC  1298 2648 A 10156 94 453 GGRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHYPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV				Į		1	VEKIIAIFGLEGLGLRDLWKPGIELLQE
TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC  1298 2648 A 10156 94 453 GGRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHYPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1297	2647	A	10138	48	407	MVSSCCGSVCSDOGCGODLCOFTCCRPSCCE
CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC  1298 2648 A 10156 94 453 GGNRKSAEMFSQVYRTPASGCYYLNSMTPEG QEMYLRFQQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIPMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKERISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1.27.	1 20 //	'''	1 .0.33	1 '		
QPVCCQPTCCRPSCCETTCCHPXCC	i .					1	
1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	ł		ì	1	
QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	ľ	1	1		1		QPVCCQPTCCRPSCCETTCCHPXCC
QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1298	2648	Α	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG
QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP   NIVAVKNDHDFLEKDLGEPLCRRLNT     1299   2649   A   10161   1   393   PRFSELVDGRGRVSARFGGSPSKAATVRSQPT   ASAQLENMEEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADAHSCDCGPGF   KGRRCELACIKVSRPCTRLFSETKAFPVWEGG   VCHHV     1300   2650   A   10162   98   391   AKIASLERIMPANYTCTRPDGDNTDFRYFIYA   VTYTGILGPGLIGNILALWVFYGYMKETKRA   VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF   VPV     1301   2651   A   10165   1   7545   PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE   KQLQCMPMEGRGRASSSISDLQGKGFEKGTG   EKHVPGVGSARHSPQASAGGSPWQRGKAQT   RWLGKPDPGRKRRGSPQEEGGLRVSAAAR   LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP   PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA   LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP   RTLSVEEPGVECNQLCLYADVTDPVLCLGQK   DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP   ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	1	1	i	1	,
INIVAVKNDHDFLEKDLGEPLCRRLNT	1	1	1	1	1	1	
1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1	1	l l	,		
ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			<u> </u>		<u> </u>	<u> </u>	
VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1299	2649	Α	10161	[ 1	393	
VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	1	1	1	1	ASAQLENMEEAPKRVSLALOLPEHGSKDIGN
KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIPMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1	1			1	
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1	J	J		1	7
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		i	1		l		
VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV					<u> </u>	<u> </u>	
VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301' 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1300	2650	A	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA
VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301' 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	1	1	1	1	VTYTGILGPGLIGNILALWVFYGYMKETKR A
1301' 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1	1	1		1	
1301' 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1	1	1	1	i	
KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	L	1	<u> </u>	<del> </del>	<del> </del>	<u> </u>	
EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1301	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE
EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	1	1	f	1	KQLQCMPMEGRGRASSSISDLOGKGFEKGTG
RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			1	1		1	
LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			1	1	1	1	
PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		}		}	j		
LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	1				
LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	1	1	ŀ		PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA
RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1		1		<b>.</b>	
DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		i	1	}	)		
ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1					•	I TATES ATTER O ATTORNATION A TOT ATTORNATION
			1		1		DDC/ACCARCEADAIGEGRAI ARADISAN
							ARRLSESLHVVDENKNESKIEREHKRRTSTPV

NO. of nucl- eoide seq- wence unnee    NO. of mucl- eoide seq- wence unnee    NO. of mucl- eoide seq- wence unnee    NO. of mucl- eoide seq- wence unnee    NO. of seq- unnee	SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
nucleotide seq- uence  uence							
Sequence					-		
uence    914   gr   gr   gr   gr   gr   gr   gr   g							
mence per de mence en		-					M-Methionine N-Assessine D-Droline
amino acid residue of peptide requence peptide sequence    Thirmonine, W-Valine, W-Typtophan, Y-Typtosin, X-Valine, W-Typtophan, Y-Typtosin, X-Valine, W-Sisp codon, /-possible nucloside deletion, wpossible nucloside deletion, wpossible nucloside deletion, wpossible nucloside deletion, wpossible nucloside deletion, wpossible nucloside deletion, wpossible nucloside deletion, wpossible nucloside deletion, wpossible nucloside insertion    XSTI, KNEKHIR, KNDDSETPHI, KSLI, KKEVKS KEKPEREKTPSEDKI, SVKHKYKGDCMRIKTO DETELHISSERGI, KVERNIKORSCO, WTL. STED, WTL. STE		uence					O-Obstania B-Assisia S-Saria
residue of peptide sequence    Poptide in unclotide delicin, "possible nuclotide delicin, "possible nuclotide delicin, "possible nuclotide delicin, "possible nuclotide insertion    RSTILKNEKHLKKDDSETPHLKSILKKEVKSS   KEKPEREKTPEDKLSVKHKYKGDCMHKITG   DETELHSSEKGLKVERINGKOSOGYKIKSDDK   TERKSKHRNERKLSVILGKDGKPVSEVIIKTDE   NVRKENNIKKERLSKILGKDGKPVSEVIIKTDE   NVRKENNIKKERLSKILGKDGKPVSEVIIKTDE   NVRKENNIKKERLSKILGKDGKPVSEVIIKTDE   NVRKENNIKKERLSKILGKDGKPVSEVIIKTDE   NVRKENNIKKERLSKILGKDGKNESSEKILKSKELVLKS   SKKTOGKOVKVPETLOBGATKQATTPKPD   REKNTEERDSBEQKRSKVENKPETETGVEPV   LETASSSAHSTOKDSSHRAKLPLAKEKYKSD   NOFSTRIE ERKLSDGKRSSKLKINSKDIKKD   REKNEDDELOGKEVDSSHRKARGNISLAMEKKL   SRICCENRROSLSQEMAKOBEKLAANTLSTP   SOSSLQRPKKSGDMTILFEQEPMEIDSEPOVE   NVFEVSKTOPINNINSIQDDSERMKOKTS   ATVOKDELRTCTADSKATAPAYKPGGTGV   NVFEVSKTOPINNINSIQDDSERMKOKTS   ATVOKDELRTCTADSKATAPAYKPGGTGV   NNSEKHADHRSTLKKMHIQSAVSKNINPGE   KEPHRIGTTEVNIDSGTVHRMLLSAPSENDRV   NNSEKHADHRSTLKKMHIQSAVSKNINPGE   KEPHRIGTTEVNIDSGTVHRMLLSAPSENDRV   NNSEKHADHRSTLVIKSESVLKTSDSKEGG   GEFTVDTPAKASTISKRHIPEAHQATLLDGKQ   GKVIMPIGSKLTGVIVESEVLKTSDSKEGG   GEFTVDTPAKASTISKRHIPEAHQATLLDGKQ   GKVIMPIGSKLTGVIVESEVLKTSDSKEGG   GEFTVDTPAKASTISKRHIPEAHQATLLDGKQ   GKVIMPIGSKLTGVIVESEVLKTSDSKEGG   GEFTVDTPAKASTISKRHIPEAHQATLLDGKQ   GKVIMPIGSKLTGVIVESEVLKTSDSKEGG   GEFTVDTPAKASTISKRHIPEAHQATLLDGKQ   GKVIMPIGSKLTGVIVESEVLKTSDSKEGG   GEFTVDTPAKASTISKRHIPEAHQATLLDGKQ   GKRENDENSEVDTSAGSGSAPSVLHQRNGGTE   VERRNENSEVDTSAGSGSAPSVLHQRNGGTE   VERRNENSEVDTSAGSGSAPSVLHQRNGGTE   VERRNENSEVDTSAGSGSAPSVLHQRNGGTE   GFASSTETLTSTKEGESGCAVAFEGRAVTE   GFASSTETLTSTKEGESGCAVAFEGRAVTE   GFASSTETLTSTKEGESGCAVAFEGRAVTE   GFASSTETLTSTKEGESGCAVAFEGRAVTE   GFASSTETLTSTKEGESGCAVAFEGRAVTE   GFASSTENHALDRAKSTERGKDEGAVTT   GFASSTETLTSTKEGESGCAVAFEGRAVTE   GFASSTENHALDRAKSTERGKDEGAVT   GFASSTENHALDRAKSTERGKDEGAVT   GFASSTENHALDRAKSTERGKDEGAVT   GFASSTENHALDRAKSTERGKDEGAVT   GFASSTENHALDRAKSTERGENDEN    DEGITTSTTGAKEGGGGGAVTSAGGEGGAVT   GFASSTENHALDRAKTAGET   GFASSTENHALDRAKSTERGENDEN    DEGITTSTTGAKEGGGGGAVTSTGGGG	uence			914			
peptide sequence    Sequence	1						
Incledide insertion	1					sequence	
RSTLKNEKHLKKDOSEPTHLKSLLKKEVKSS  KKEPERKTYSEOKLSVKHKYKGOCHMIKTO DETELHSSEKGLKVERNIQKOSQOTKLSSIDDK TERKSKRINBERLSLYGKORGKYVSEYIKTDE NVRKENNIKKERRI SAEKTKAEHKSRSSIDSK IQKOSLGSKQHGHTLORGKSSYSEDKCOMDST NMDSNLKPEEVVHKEKRIKSLEKLVLKS SKYGGLGVVEVVETLEDGAATKQATTKEPD KEKNTEENISERQKKSKVEDKFFEETGVEPV LETASSSAHSTQKDSSHKARGKSKSLEDKKKLD BENDELOGGEVEVSERSKARGKSKUSDKYFEETGVEPV LETASSSAHSTQKDSSHKARGNSSLEMEKKL SKRTGERDSEKGRKSKVEDKFFEETGVEPV LETASSSAHSTQKDSSHKARGNSLSMEKKL SKRTCENRIGGIS QEMAKGERELAANTLSTP SOSSLGRPKKSGDUTLIPEGEPAMEDSHEKKK SKRTCENRIGGIS QEMAKGERELAANTLSTP SOSSLGRPKKSGDUTLIPEGEPAMEDSHEKKK SKRTCENRIGGIS QEMAKGERELAANTLSTP SOSSLGRPKKSGDUTLIPEGEPAMEDSHERWIK SKRLEDHERTSTLTKEMHIGSAVSKMINDEE EKPHIEGTTEVNIDSETVHEMLLSAPSKORGTOV NSNSEKHADHESTLTKKMHIGSAVSKMINDEE KEPHIEGTTEVNIDSETVHEMLLSAPSKORGTOV NSNSEKHADHESTLTKKMHIGSAVSKMINDEE KEPHIEGTTEVNIDSETVHEMLLSAPSKORGTOV NSNSEKHADHESTLTKKSMHIGSAVSKMINDEE KEPHIEGTTEVNIDSETVHEMLLSAPSKORGTOV NSNSEKHADHESTLTKKSMHIGSAVSKMINDEE KEPHIEGTTEVNIDSETVHEMLLSAPSKORGTOV NSNSEKHADHESTLTKKSMHIGSAVSKMINDEE KEPHIEGTTEVNIDSETVHEMLLSAPSKUNDISSE LSSVTVYVFLRESVDPVIFLDKRYTVLGGSTA STSPADHSALPNQSLTVERSEPLKTSDSKEGG GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTPAKASTISKREHEPAADATLLDGKQ GEFTVDTFACKASTISKREBEDGAAVTE GARSTETLITSKEGESGGCAVASSEDRAADATL VARAVKISANNIVENCOMATSTER TORGRESSEVDISTAGGGNAMRAMGRKSTETG TYTCTGABGRSINEVICSVTGAGPREEMAVT GGRATISTICHEDGGGAASCTGSEDSSEGPAHS SEEDGGGNANSTOWTSTGRQSGIGTVUH VERABAGAMMANENNVDSNIGTEKGSKDT GGSAVTSTGTGEGGGAAVTGAGGCTAVGE GGRATSTGTTEGGESGCAVTGAGGGTAVTGAGGE GTTASASOQSDSQLEKVEDTISTGLYGGG TYLVXGGVPECEVAHTSFGCKGCAVTGG GGRATSTGTTEGGESGCALISTISTGECECASVS GVVVSENBRAAGTVUSTAGGGGISTSSVEDC  GGRATSTGTTEGGESGCALISTISTGECECASVS GVVVSENBRAAGTVUSSTGAGGSGUSTSSVEDC  GGRATSASOQSOLEKVEDTISTGLYGGGG GGPTVAAVSEEGFFILPSSATTIKCAES LQPVAAAVEERTATGV	1	,		1			
KEKPEREKTPEEDKLSVAHKYKGDCAMIKITG DETELIASSEKGLKVERMIQKSQGYKLSSIDK TERKSKHRINERLSVIGKDGKPVSEVITKITDE NVRKERNIKKERRIASSEKTKABHIKSRSSIDSK JQKDSI.GSKQHOTTLORRSISSYSEDKCDMDST NADSSIL KPEEVVJHEKRRITSKILLEKLVLKS KSKTOGK.QVKVVETELORGATKOATTPKEPD KEKNTEENDSEKQKEKSKRTTSKILLEKLVLKS KSKTOGK.QVKVVETELORGATKOATTPKEPD KEKNTEENDSEKQKEKSKRTSKSLLEKLVLKS KSKTOGK.QVKVVETELORGATKOATTPKEPD KEKNTEENDSEKQKEKSKVEDKFETGVEPV LETASSAAHSTQKDSSHRAKKLRAKTKSKE BYSTSTILLERLSIGHKSRSLKKKK BYSTSTILLERLSIGHKSRSLKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERLSIGHKSRSLKKKKK BYSTSTILLERGERALDSERGYBK KOMPANIEN SOSSLARFKKSGDMTLITEGERALDSERGYBK NYFEVSKTODDRINNNSHQDIDSEMMKQKTS ATVQCODELRTCTADSKATAPAYKRGGTGV NSNSEKHADHRSTLTKKMHIGSA VSKMNPGE KEPHRGTTEVINDSETVHRMLLASPENDRV QKNLKNTAAEBHVAQGDATLEHSTNLDSSESG GEFTVOTPTAKASTISKRHPEAHQATLLDGKQ GKVIMPLGSKLTGVIVERESVIKTSDSKEGG GEFTVOTPTAKASTISKRHPEAHQATLLDGKQ GKVIMPLGSKLTGVIVERESVIKTSDSKEGG GEFTVOTPTAKASTISKRHPEAHQATLLDGKQ GKVIMPLGSKLTGVIVERESVIKTSDSKEGG VERRENSEVDTSAGGSASPVLHORNOGTE DVATOPRRAEKTSVATSTEGKNDAVTLSFVK AGPATTTSSETRGSEVALPCTISEADGLIGT HSRNNPLHVGABASECTVFAAAEGGGAVVTE DVATOPRRAEKTSVATSTEGKNDVTLSFVK AGPATTTSSETRGSEVALPCTISEADGLIGT HSRNNPLHVGABASECTVFAAAEGGGAVVTLSFVK GABSETFILTSTKEGEGEGCAVAESGRAADL LAVHAVKIEANVNSVYEEKDDAVTSAGSEE KODGSLASSBEVEGEGGAVANGOTE DVATOPRRAEKTSVATSTEGKNDKDVTLSFVK AGPATTTSSETRGSEVALPCTISEADGAVTAG TERRAGSISSEEVOGGONMMRMOPKETEG GESAAVTSTGTGEGGEGAVTSGEGGGSVANDOTE GESAATTSTGTEGGEGAPATGCSSTGTSCOGGANG SESSEMGEAMDSTVAKEGTNVPLVAAGCCO DEGIVTSTAAEGGEGGAVTOTE GESAATTSTGTEGGEGAPASTGGGGSTAGGGGGANTDOTE GESAATTSTGTEGGEGAPASTGGGGGSATDAGGGGGGGANTDOTE GESAATTSTGTEGGEGAPASTGGGGGSATTAGATT SESSEENGESAMDSTVAKEGGTNVPLVAAGCCC GROTSAAAGGGGGGSVTNDOTE GESAATTSTGTEGGEGAPASTGGGGGGANTDOTE GESAATTSTGAAEGGGGGSTANDSTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG					sequence		
DETELHSSERGILKVEENIGKQSQOTKLSSDSSSS TORKSENSKHENREKLSVIGKGDGFVSEVIIKTDE NVEKENNIKKERRLSAEKTKAEHEKSRRSSDSS (IQKDSSLGSKOHGTIO, QRESEYSEDKCDMDST NMDSNLKPEVVHKERRRITKSLLEEKLVLKS KSKTQGKQVKVVEPTELQGGATKQATTEKPD KEENTEENDEKGRISKVEDKTPETGVEPV LETASSSAHSTGKOSSIRAKLIAEKTKACHTKKPD KEENTEENDEKGRISKVEDKTPETGVEPV LETASSSAHSTGKOSSIRAKLIAEKTKACHTKKPD KEENTEENDEKGRISKVEDKTPETGVEPV LETASSSAHSTGKOSSIRAKLIAEKTKACHTKKPD KEENTEENDEKGRISKORISKVEDKTPETGVEPV LETASSSAHSTGKOSSIRAKLIAESKOITKKKL SRILCENREGIS, QEMAKGEKLAANTLISK DENKEKL SRILCENREGIS, QEMAKGEKLAANTLISK SRILCENREGIS, QEMAKGEKLAANTLISK SRILCENREGIS, QEMAKGEKLAANTLISK NOPEVSKTQDNRINNISHQDDSENMKQKTS ATVQKDELRTCTADSKATDAPKGRTGV NSNSKHALDHESTLIKKHHIQSAVSKMINGE KEPHERGTTEVNIDESTVERMLELAESTA ATVQKDELRTCTADSKATDAPKGRTGV NSNSKHALDHESTLIKKHHIQSAVSKMINGE KEPHERGTTEVNIDESTVERMLELAGSTA ASTSPADHSALPNQSLTVRESEVILKTSDSKGG GEGTVDTIPAKASTIKKHIPBAHGATLLDGSTA STSPADHSALPNQSLTVRESEVILKTSDSKGG GEGTVDTIPAKASTIKKHIPBAHGATLLDGSTA STSPADHSALPNQSLTVRESEVILKTSDSKGG GEGTVDTIPAKASTIKKHIPBAHGATLLDGSTA KKENDINAEPHLIAGSTA STREAKTORATLLOSKOG GEGTVDTIPAKASTIKKHIPBAHGATLLDGSTA STSPADHSALPNQSLTVRESEVILKTSDSKGG GEGTVDTIPAKASTIKKHIPBAHGATLLDGSTA KKENDINAEPHLIAGSTANTKHPBAHGATLLDGSTA STSPADHSALPNQSLTVRESEVILKTSDSKGG GEGTVDTAPAKASTIKKHPBAHGATLLDGSTA KKENDINAEPHLIAGSTANTKHPBAHGATLLDGSTA KKENDINAEPHLIAGSTANTKHPBAHGATLLDGSTA KKENDINAEPHLIAGSTANTKHPBAHGATLLDGSTA KKENDINAEPHLIAGSTANTKHPBAHGATLLDGSTA KKENDINAEPHLIAGSTANTKHPBAHGATLLDGSTANTKHPBAHGATLANTKHPBAHGATLLDGSTANTKHPBAHGATLLDGSTANTKHPBAHGATLLDGSTANTKHPBAHG							KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
TERKSKHINERKI.SVJGKOGKPYSEVIIKT)  NYKKENNIKKERRISABEKTARHEKSRSSDSK 10KDSLGSKOHGITLORRSESYSEDKCUMDIST NIMDSNI.KPIEVVHKERRITKISLLEEKI.VLKS KSKTQGKQWKVVETELQBGATKQATTPKIP KEKMTEENDESKORKSKVEDEPRETGVEPV KEKMTEENDESKORKSKVEDEPRETGVEPV LETASSSAHSTOKDSSHRAKI.PLAKERKYKSL KOSTSTRI.ERKI.SDOHFKSSLLIKHSKDIKKKD ENKSDDK.DOKEVDSSHEKARGNSSLIMEKK. SIRLCENRIGGI.SQDMAKGEEKI.AAVILSTJP SOSSLQRRKKSGDMTI.PEOEPMEIDSEPGVEP NYFEVSKTQDRINNISHQDDSENMKOKTS ATVQKDEILTI.TADSKATAPAYRGRGTOV NSNSSKHALDIRSTI.TKKMIIQSAVSKMIPGE KEPHROTTEVNIDSTURMILSSTVERMILSSAVSKAMPGE KEPHROTTEVNIDSTVERMILSSAVSKAMPGE KEPHROTTEVNIDSTVERMILSSAVSKAMPGE KEPHROTTEVNIDSTVERMILSSAVSKAMPGE GEFTOTTPAKASTIKKHPEAHQATILDIGKQ GKVIMPLOSKLITOVIENENITKEGGI.VDM KKENDLNAEPNILAGTIKATVEKKENGIAVD HVVGLNTEKYAETVKLIKHRSPGKVKDISID VERRYENSEVDTSAGSSASPSVHLORNOGOTE DVATGPRAEKTSVATSTECKDKDVTLSYVK AGPATTTSSETRGSVALPCTSIELADGGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVSERAADL LAVHAVKLEANVINSVYTEEKDDAYTSAGS KCGSTTTSSTAGSSAPSVHLORNOGOTE DVATGPRAEKTSVATSTECKDROYTSAVS AGPATTSSETRGSSVALPCTSIELADGGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVSETE GFAESETFLTSTKEGESGECAVSETE GFAESETFLTSTKEGESGECAVSTEG TERAGSISSEEVDGSOGNMMRMGPKKETTE GFAESETFLTSTKEGESGECAVSTEG TERAGSISSEEVDGSOGNMAMMGPKKETTE GESAVTSTGTTEGEGGGASVNTOGTE  DSAVTSTGTAGBGSBNYTVESTGAGREEMMY GAGVVLGDNDAPPGTSASQEGGSSVALPCTSTEG TTCTGABGRSDNYTVESVTGAGREEMMY GAGVVLGDNDAPPGTSASQEGGSSVALPCTSTEG TERAGSISSEEVDGGGASVNTOGTE  DEGINTSTGAKEGEGGASVTTOSTEVSTORGREEM TOCTGABGRSDNYTVESVTGAGREEMMY GAGVVLGDNDAPPGTSASQEGDGSVNTOTE  GESAVTSTGTTEDEGGBACTGTSTGGNEGGAP SESEENGESAMDSTVAKEGTNYPLVAAGEGGAVVL DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDGLEKVEDTTSTGLVGOS TOLLVSGEVPECSVAHTSPEKEDDBLITTST. BEGCEAVMIGAVLAGDTANSPEKEDDDLTSTGARGEGG PMTSAASDQSDGLEKVEDTTSTGGLOSG TOLLVSGEVPECSVAHTSPEKEDDDLTSTGARGEGG PMTSAASDGSDSLEKVEDTTSTGLVGOS GVVVSENBERGGTAVMEEKGGGGTSVNTOGTE BNMEGTTVTTEEFFAMPSAVGDDSGLAKTETTSTGLVGOS GVVVSENBERGGTVMEEKGGGGGTSVNTOGT TERAGSISSTAADAMPSAVGROPSVTAREGGGGPTANSTA	1					;	KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
TERKSKHINERKI.SVJGKOGKPYSEVIIKT)  NYKKENNIKKERRISABEKTARHEKSRSSDSK 10KDSLGSKOHGITLORRSESYSEDKCUMDIST NIMDSNI.KPIEVVHKERRITKISLLEEKI.VLKS KSKTQGKQWKVVETELQBGATKQATTPKIP KEKMTEENDESKORKSKVEDEPRETGVEPV KEKMTEENDESKORKSKVEDEPRETGVEPV LETASSSAHSTOKDSSHRAKI.PLAKERKYKSL KOSTSTRI.ERKI.SDOHFKSSLLIKHSKDIKKKD ENKSDDK.DOKEVDSSHEKARGNSSLIMEKK. SIRLCENRIGGI.SQDMAKGEEKI.AAVILSTJP SOSSLQRRKKSGDMTI.PEOEPMEIDSEPGVEP NYFEVSKTQDRINNISHQDDSENMKOKTS ATVQKDEILTI.TADSKATAPAYRGRGTOV NSNSSKHALDIRSTI.TKKMIIQSAVSKMIPGE KEPHROTTEVNIDSTURMILSSTVERMILSSAVSKAMPGE KEPHROTTEVNIDSTVERMILSSAVSKAMPGE KEPHROTTEVNIDSTVERMILSSAVSKAMPGE KEPHROTTEVNIDSTVERMILSSAVSKAMPGE GEFTOTTPAKASTIKKHPEAHQATILDIGKQ GKVIMPLOSKLITOVIENENITKEGGI.VDM KKENDLNAEPNILAGTIKATVEKKENGIAVD HVVGLNTEKYAETVKLIKHRSPGKVKDISID VERRYENSEVDTSAGSSASPSVHLORNOGOTE DVATGPRAEKTSVATSTECKDKDVTLSYVK AGPATTTSSETRGSVALPCTSIELADGGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVSERAADL LAVHAVKLEANVINSVYTEEKDDAYTSAGS KCGSTTTSSTAGSSAPSVHLORNOGOTE DVATGPRAEKTSVATSTECKDROYTSAVS AGPATTSSETRGSSVALPCTSIELADGGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVSETE GFAESETFLTSTKEGESGECAVSETE GFAESETFLTSTKEGESGECAVSTEG TERAGSISSEEVDGSOGNMMRMGPKKETTE GFAESETFLTSTKEGESGECAVSTEG TERAGSISSEEVDGSOGNMAMMGPKKETTE GESAVTSTGTTEGEGGGASVNTOGTE  DSAVTSTGTAGBGSBNYTVESTGAGREEMMY GAGVVLGDNDAPPGTSASQEGGSSVALPCTSTEG TTCTGABGRSDNYTVESVTGAGREEMMY GAGVVLGDNDAPPGTSASQEGGSSVALPCTSTEG TERAGSISSEEVDGGGASVNTOGTE  DEGINTSTGAKEGEGGASVTTOSTEVSTORGREEM TOCTGABGRSDNYTVESVTGAGREEMMY GAGVVLGDNDAPPGTSASQEGDGSVNTOTE  GESAVTSTGTTEDEGGBACTGTSTGGNEGGAP SESEENGESAMDSTVAKEGTNYPLVAAGEGGAVVL DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDGLEKVEDTTSTGLVGOS TOLLVSGEVPECSVAHTSPEKEDDBLITTST. BEGCEAVMIGAVLAGDTANSPEKEDDDLTSTGARGEGG PMTSAASDQSDGLEKVEDTTSTGGLOSG TOLLVSGEVPECSVAHTSPEKEDDDLTSTGARGEGG PMTSAASDGSDSLEKVEDTTSTGLVGOS GVVVSENBERGGTAVMEEKGGGGTSVNTOGTE BNMEGTTVTTEEFFAMPSAVGDDSGLAKTETTSTGLVGOS GVVVSENBERGGTVMEEKGGGGGTSVNTOGT TERAGSISSTAADAMPSAVGROPSVTAREGGGGPTANSTA	ŀ						DETELHSSEKGLKVEENIOKOSOOTKLSSDDK
NVKKENNKKERRLSAEKTKAEHKSRRSSISS  IQKOSIOLGSKOHGTIQ.QRSESYSEDKCDMDST  NMDSNLKPEVVHKERRRTKSLLEEKL.VLK  KISKTOGKQVKVVPTEI.QBGATKQATTPKPD  KEKNTEENDSEKORKSKVEDKPFEETGVEPV  LETASSAHTOKOSSHRAKULPKGATTPKPD  KEKNTEENDSEKORKSKVEDKPFEETGVEPV  LETASSAHTOKOSSHRAKULPKGATTPKPD  KEKNTEENDSEKORKSKVEDKPFEETGVEPV  LETASSAHTOKOSSHRAKULPKGATTPKVE  KEKNTEENDSEKORKSKVEDKPFEETGVEPV  LETASSAHTOKOSSHRAKULPKGATTPKVE  KOSTSTRLERKLSDCHKSRSLKHSSKDIKKKD  ENKSDDKDOKEVDSSHEARANSSLAEKKYKSD  KOSTSTRLERKLSDCHKSRSLKHSSKDIKKKD  ENKSDDKDOKEVDSSHEARANSSLAEKKYKSD  KOSTSTRLERKLSDCHKSRSLKHSSKDIKKKD  SOSSLQRPKKSGDMTLIFEOPMEIDSETGVIKTST  SOSSLQRPKKSGDMTLIFEOPMEIDSETGVIKTST  SOSSLQRPKKSGDMTLIFEOPMEIDSETGVIKTST  NEVERVANGENGE  KEPHRGTITEVNIDSETGVIKTMLSAFSENDRV  VERNENSETLIKKMHIQSA VSKMNFOE  KEPHRGTITEVNIDSETGVIKTMLSAFSENDRV  VERNENSETLIKKMHIQSA VSKMNFOE  KEPHRGTITEVNIDSETGVIKTMLSAFSENDRV  VERNENSETLIKKMHIQSA VSKMNFOE  KEPHRGTITEVNIDSETGVIKTMLSAFSENDRV  VERNENSETTISKHIPBAHQATILLOST  STSPADHSALPNOSITVRESEVIKTSSKEGG  EGFTVDTPAKASITSKHIPBAHQATILLOST  STSPADHSALPNOSITVRESEVIKTSSKEGG  EGFTVDTPAKASITSKHIPBAHQATILLOST  VERNENSEVITVRESTYRTSKERKDELOST  VERNENSEVITVRESTYRTSKERKDELOST  VERNENSEVITVRESTYRTSKERKDELOST  VERNENSEVITVRESTYRTSKERKDELOST  VERNENSEVITVRESTYRTSKERKDELOST  VERNENSEVITVRESTYRTSKERKDELOST  DVATGPRAEKTSVATSTECKDKDVILST  VERNENSEVITTSTSTRQSEVALPCTSIRADEGLIGT  HSENNPH HVGAGASECTVALAAEGGAVVTE  GFAESETILTSTKEGGSGCAVAESEDRAADL  LAVHAYKLBANVRVVTEEKDADAVTSAGSES  KCDGSLSRDSEIVEGTTITTISEVESDGAVTSAG  TEIRAGSISSEVDGSQGNMMMMGPKKETEG  TEIRAGSISSEVDGSQGNMMMMGPKKETEG  TEIRAGSISSEVDGSQGNMMMMGPKKETEG  TEIRAGSISSEVDGSQGNMMMMGPKKETEG  TEIRAGSISSEVDGSQGNMMMMGPKKETEG  TEIRAGSISSEVDGSQGNMMMMGPKKETEG  GESAVTSGTGTEGGFASCTOSEDSSEGFAIS  SESEENGESAMDSTVAKEGTNVPLVAAGGC  DEGITYTSTGAKEGEDEGDEVTSTGRGNEGH  ASTCTGLGESEGVLICESAEGDSSGGTAVDE  GESENGESAMDSTVAKEGTNVPLVAAGGC  DEGITYTSTGAKEGEDEGEGEVTSTGRGNEGH  ASTCTGLGESEGVLICESAEGDSGGGTAVDE  VERLOGAMTTASAGDINNSHALHRAPPTSGGLGGAAALEE  LOPVAAAVEERANDSNDSMSTGEGSGGTANSTGGLESCOOL  TOSSAKGTUSSTAAAVNRANDMSMSGDGSGLSATEG  GDLSATEVSKHKVPMPSLLAENG	ļ		1			•	
IQKDSLGSKOHGITLQRESEYSEDKCMDDS  NMDSNLKPEVVHERRITKSLEEKLVLKS  KSKTOGKQVKVVTFIELGEATKQATTFKPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKQKSKVEDEPFETOVEPP  KEKNTEENDSEKGKSKVEDEPFETOVEPP  KEKNTEENDSEKGKSKVEDEPETOVEPP  KOPTITELSSENDERGKENDSEKGKSKUEDEPSETOVE  KOPTITELSSENDERGKENDSEKGKENSSLHERKL  SRRLCENRRGSLSGEMARGEKLAANTLIST  SGSLQAPKKSGIDMTLIPEGPFENDSEFOVE  NVFEVSKTODNTRINSHQDIDSENMGKITS  ATVQKDELTICTADSKATAPAYRGRGTGV  NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE  KEPHIRGTTEVNIDSETVIRMLISTSENDSESSES  LSSVTVPLRESPYDDVIELFBKTVLCGSTS  GKINKINTAAEBHVAQGDATLEHISTINLDSSPS  LSSVTVPLRESPYDDVIELFBKTVLCGSTS  STSPADHSALPNQSLTVRESEVLKTSDSKEGG  GGFTVDTPAKASTIKKHIPBAHQTLLDGKQ  GKVIMPLOSKLTOVIVERENTIKEGGI VDMA  KKENDLNAEPHLKQGTKATVENGKKDGIAVA  KKENDLNAEPHLKQGTKATVENGKKDGIAVA  HVVGLNTEKY AFTVKLKHRRSPGKVKDISID  VERRNENSEVDTSAGGSGAPSVLHORNOGTIT  DVATGPRRAEKTSVATSTEKKDKVTLSFVK  AGPATITSSETRQSEVALPCTSIGADEGLIGIT  HSRNNPLHVGGASSECTVFAAABGGGVVTE  GRAESETFLTSTKEGESGCAVAESEDRAADL  LAVHAVKLEANVSVVTEEKDDAVTSAGGSEE  KCDGSLSRBGEVGATITISEVESDGAVTSAG  GRIGNINSESVENGVAGFREEMOTT  GRAESETFLTSTKGGESGGCAVAESEDRAADL  LAVHAVKLEANVSVVTEEKDDAVTSAGGSEE  KCDGSLSRBSEIVGGTITISEVESDGAVTSAG  TEIRAGSISSEEVDAGFTSAGGSTSGGSENSOFTE  GESAVTSTGITEDGEGPASCTGGEBSSEGFARS  SESEENGGSAMDSTVAKSTOTAVGAFERGSKDT  GRAVTSTGITEDGEGPASCTGGEBSSEGFARS  SESEENGGSAMDSTVAKSTOTAVGTERGKDTOT  GRAVTSTGITEDGEGPASCTGGSEDSSGFARS  SESEENGGSAMDSTVAKSTOTAVGTERGKDTOT  GRAVTSTGITEDGEGPASCTGGSEDSSGFARS  SESEENGGSAMDSTVAKSTOTAVGTERGKDTOT  GRAVTSTGITEDGEGPASCTGGSEDSSGFARS  SESEENGGSAMDSTVAKSTOTAVGTERGKDTOT  GRAVTSTGITEDGEGPASCTGGSEDSSGFARS  SESEENGGSAMDSTVAKSTOTAVGTERGKDTOT  GRAVTSTGITAGAGETERGKDT  DICSSAKGTVESSYTSAVSGRDEVTPYTGGG  GPMTSAASDGSSDGLEKVEDTTISTGLVGGG  GPMTSTAASDGSSDGLEKVEDTTISTGLVGGG  GPMTSAASDGSSDGLEKVEDTTISTGLVGGG  GPWTSAAVDGEGDPSVTTAEBMGDTAMSTSVEDG  GOLSATEVSKHKVPMPSLLARNCRCFQPYP  GGEFFQPVLAVSTERGEMOPSYSTIKFSAQGGH  PSAVCAEKERKHKECPELOPFAGRGGE	i						
NMDSNLKPÉVVIKÉRRATKSLLEEIKJUKÉ KSKTOGKOVKVETTELQGATTKAP KEKNTEENDEKORKSKVEDKPFETTGVEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP LETASSAHSTOKDSSHRAKUL AKEKYKEP SOSSLQRAKGEKLAANTLSTSP LSSVTVVPLRESYDPDVPLPDKRTVLEGSTA STSPADISALPYOKJUTENGATTSSKRIPGAHQATLLDOK GKUMPLOSKLTOTVENENTIKEGGLVDMA KKENDLNAEPHLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHRKSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNOGTE DVATGPRAEKTSVATSTEGKDKJOND VERRIENSEVDTSAGSGSAPSVLHQRNOGTE DVATGPRAEKTSVATSTEGKDKADAT HVVGLNTEKYAETVKLKHRKSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNOGTE DVATGPRAEKTSVATSTEGKDKADATS AGPATTTSSTRQSEVALPCTSIEADEGLIGIT HSRNIPH HVGAEASECTVAAAEGGAAVVTE GFAESETILTSKEGGSGECAVAESEDRAADL LAVHAYKLEANVNSVVTEEKDADAVTSAGSEE KCDGSLSRDSEIVEGTTITISEVESDQAVTSAG TEIRAGSISSEVDGSQGMMMMMGPKKETE TICTGAEGESGOVATORFT GERSEL TOTCTGAEGEGGDVNTORTFT GERSEL TOTCTGAEGEGGDVNTORTFT GERSEL GGAVVLGDINDAPPOTSASQEGDGSVNDGTFT GGSAVTSTOTTGEGGPASCTOSEDSSEGFAIS SESEENGGSAMDSTVAKEGTNVFLVAAGGCE GPMTTSAASDQSDSQLEKVEDTITTSTGLVGGGE GPMTTSAASDQSDSQLEKVEDTITTSTGLVGGGE GPMTTSAASDQSDSQLEKVEDTITTSTGLVGGGE GPMTTSAASDQSDSQLEKVEDTITTSTGLVGGGE GPMTSAASDQSDSQLEKVEDTITTSTGLVGGGE GPMTSAASDQSDSQLEKVEDTITTSTGLVGGGE GPMTSAASDQSDGLEKVEDTITTSTGLVGGGE GPMTSAASDGSDGLEKVEDTITTSTGLVGGGE GPMTSAASDGSDGLEKVEDTITTSTGLVGGGE GPMTSAASDGSDGLEKVEDTITTSTGLVGGGE GPMTSAASDGSDGLEKVEDTITTSTGLVGGGE GPMTSAASDGRANGSPALHAGRIGDGLAAGGSTOMA ALISTSTRAECMPISAGDGHAAGGCCPOPVR GGEGPOPLANTSAGDGGLAAGGSTAAAB AL	1	}	<b>[</b>				
KSKTOGKOVKVYETELGEGATKOATTYRED KEKNTEENDSERQKESSEVEDKPEETGVEPV LETASSSAHSTQKDSSHRAKI, PLAKEKYKSD KDSTSTRILERILSDGHKSSIKHSSKDIKKKD ENKSDDKDGKEVDSSHRAKI, PLAKEKYKSD KDSTSTRILERILSDGHKSSIKHSSKDIKKKD ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRALCENBRGSIS, GEBAKGEBKLAANTISTP SGSSIQRYKKSGDMTLIPEGEBUSERGVE NVFEVSKTQDINENNISHQDIDSENMKGKTS ATVQKDELRICTADSKATAPAYRFGRGTGV NSNSEKHADHRSTLIKKMHIQSAVSKMNFGE KEPIHRGTTEVNIDSETVHRMLISAPSENDRY QKNILKNTAAEEHVAQGDATLEHISTILDSSPS SSTYVPLRESTYDPDVIPLEDKRTVLEGSTA STSPADHSALPNQSLIVRRESEVLKTSDSKEGG GEFTVDTPAKASTISKRHPEARQATLLDSSPS LSSVTVVPLRESTYDPDVIPLEDKRTVLEGSTA STSPADHSALPNQSLIVRRESEVLKTSDSKEGG GEFTVDTPAKASTISKRHPEARQATLLDSSPS GEFTVDTPAKASTISKRHPEARQATLLDSSPS USSTVEVLRESTYDPDVIPLEDKRTVLEGSTA STSPADHSALPNQSLIVRRESEVLKTSDSKEGG GEFTVDTPAKASTISKRHPEARQATLLDSSPS LSSVTVPLRESTYDPDVIPLEDKRTVLEGSTA STSPADHSALPNQSLIVRRESEVLKTSDSKEGG GEFTVDTPAKASTISKRHPEARQATLLDSSPS LSSVTVPLRESTYDPDVIPLARQATLLDSSPS LSSVTVPLRESTYDPDVIPLARQATLLDSSPS LSSVTVPLRESTYDPOVIPLARQATLLDSSPS LSSVTVPLRESTYDPOVIPLARQATLLDSSPS LSSVTVPLRESTYDPOVIPLARQATLLDSSPS LSSVTVPLRESTYDDSTVERSTYLLESTYLLDSSPS LSSVTVPLRESTYDDSTVERSTYLLESTYLLDSSPS LSSVTVPLARGATLLDSSPS LSSVTVPLRESTYDDSTVERSTYLLGSSPS LSSVTVPLARGATLLDSSPS LSSVTVPLARGATLLDSSPS LSSVTVPLARGATLLDSSPS LSSVTVPLARGATLLSSPSLARGATLLSSPSLARGATLLDSSPSLARGATLLSSPSLARGATLLDSSPSLARGATLLSSPSLARGATLLDSSPSLARGATLLSSPSLARGATLLDSSPSLARGATLLSSPSLARGATLLSSPSLARGATLLSSPSLARGATLLSSPSLARGATLLSSPSLARGATLSSPS	}	ĺ					l •
KEKNTEENDSEKQRKSKVEUKPFETTGVEPV LETASSAAHTSCROSHRAKLPLAKERYKSD KDSTSTRIERKLSDGHKSRSLKHSKOINEKAL ENKSDDKDGKEVEDVSSHEKARAGNSSLMEKKL SRALCENRRGSLSQEMAKGEKLAANTLSTP SGSSLQRPKKSGDMTLIFEQEPMEIDSERGYE NVEVSKTQDINENNINSHQDIDSENMEQKTS ATVQKDELRTCTADSKATAPAYRPGGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE KEPHRGTTEVNIDSETVHRMILSAPSENDRV QKNILKNTAABEHVAQGDATENITLSSPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	1	1	<b>}</b> '	ł	į		
LETASSAHSTQKDSHKRILHSSKDIKKKU KDSTSTRILERILSDGKKSRILHSSKDIKKKU ENKSDDKDQKEVDSSHEKARANSSLMEKKI. SRRLCENRRGSLSQEMAKGEKLAANTLSTP SGSSLQRKKSGDMTLIFEQEPMEIDSEPGVE NVFEVSKTQDINRNNSHQDIDSENMKQKTS ATVQKDELRICTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMPGE KEPIFRGTTEVNIDSETVTRMLISAPSENDRV QKNLKNTAAEEHVAQGDATLEHISTNLDSSPS LSSVTVVPLRSTVDPVPUFLDSSTS LSSVTVVPLRSSTVDPVPUFLDSSTS LSSVTVPLRSSTVDPVPUFLDSSTS LSSVTVPLRSSTVDPVPUFLDSSTS ASTSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRIIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENSTRIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGAVD HVVGLNTEKYAETVKLKIKKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHQRNGGTE DVATGPRRAEKTSVATSTEGLOKDVTLSPYK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGLAASASCTVFAEEGGAVVTE GFAESETFLTSTKEGESGGCAVAESEDRAADL LAVHAVKIEANNISVTEEKDDAVTSAGSEE KCDGSLSRDSSEVEGTITTISTEVSEDGAAVTSAG TEIRAGSISSEEVDGSQGNMMRMOFKKETG TVTCTGAEGRSINPVLCSVTGAGPREERMVT GAGVVLGDNDAPPGTSASGEGGSVSNIGGTE GESAVTSTGITEDGEGPASCTGGSPSNIGGTE GESAVTSTGITEDGEGPASCTGGSPSNIGGTE GESAVTSTGITEDGEGPASCTGGSPSNIGGTE GESAVTSTGITEDGEGPASCTGTTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQGTVVEH VEAEAGAAIMANNENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKGEDVTFVPGGCE GPMTSAASDQSDSQLEKVEDTTTSTGCVGGS YDVLVSGEVPEGCVAHTSPSFEKDEDUITSVE NEECDGLMATTAGSGDTINQNSLAGGKNQK VLISTSTTNDVTTQVSAITDVFGGGLS GEPMSAASDQSDSQLEKVEDTTTSTGCABEIGH ASTCTGLGEESEGVLICESAEGDSGGALSSE LQEVAAAVEGRATASGDTINQNSLAGGKOGK VLISTSTTNDVTTQVSAITDVFGGGLSDARRTE ENMEGTRYTTEEFFEAPMPSAVSGDDSQLTAS RSEECDGCAMTTASGDTINQNSLAGGKOGK VLISTSTTNDVTTQVSAITDVFGGGLSDARRTE ENMEGTRYTTEEFFEAPMPSAVSGDDSQLTAS RSEECDGCAMTTSTSIGEEFFELPISATTIKCAES LQEVAAAVEGRATATSGEDTITSTVEDGGS GOLSSAKGIVEGAVTATORGREIGH ASTCTGLGEEGGVTTTTSTGATGTAGGST TSGCCAVMIGGAVLQDEDGTTTTTRVEDLSDA AIISTSTTAECMPSAASDREEPEPPPTAGGGGGSTGTAAGGSST TSGCCAVMIGGAVLQDEDGTTTTRVEDLSDA AIISTSTAECMPSAASDREEPEPPPTAGGGGGSTGTAAGGSST TSGCCAVMIGGAVLQDEDGTTTTRVEDLSDA AIISTSTAECMPSAASDREPEPPPTAGGGGSTGTAAGGSST TSGCCAVMIGGAVLQDEDGTAADMPTVQCTTTTAGGSTGTAAG	1	1	1				
KDSTSTRLERKLSDGHKSRSLKHSKNICKKD ENKSDDKDGKEVDSSHEKARRONSSIMEKKL SRICLERURGGIS OGMAKGEKLAANTLSTP SGSSLQRPKKSGDWTLIPEQEPWEIDSERGYE NVFEYSKTQDNRNINNSRQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYREGGTOV NSNSEKHADHRSTLTKKMHUGSAVSKMNPGE KEPHRGTTEVINDSETVHRMLLSAPSEMDRV QKNLKNTAAERIVACQDATCHENSTILDSSPS LSSVTVVPLRESVDPDVPLFDKTVLLGSTA STSPADHSALPNQSLTVRESVLKTSDSKEGG GEFTVDTPAKASITSKRHIPEARQATLLDGKQ GKVIMPLOSKLTGVENENSTEGGLVDMA KKENDLNAEPHLKQTTKATVENCKKDGLAVD HVVGLNTEKYAETVSKLKTSDSKEGG GEGTTVDTPAKASITSKRHIPEARQATLLDGKQ GKVIMPLOSKLTGVENENSTEGGLVDMA KKENDLNAEPHLKQTTKATVENCKKDGLAVD HVVGLNTEKYAETVSKLKHKRSPGKVKDISID VERRINENSEVDTSAGSGAPSVLHQRNQQTE DVATGRRAEKTSVATSTEGKDKDATLSPVK AGPATITSSETRGSEVALPCTBADGLIGT HSRNIPLHVGABASECTVFAAAEEGAAVTE GFAESETHLTSTKEGESGEACHAEBGALGIGT HSRNIPLHVGABASECTVFAAAEEGAAVTE GFAESETHLTSTKEGESGEACHAEBGALGATSAG TERAGSISSEEVDGSQONMARMOPKKETEG TVTCTGAEGRSDIVPGITTIFSEVSBDGAVTSAG TERAGSISSEEVDGSQONMARMOPKKETEG TVTCTGAEGRSDIVPGITTIFSEVSBDGAVTSAG TERAGSISSEEVDGSQONMARMOPKKETEG TVTCTGAEGRSDIVPGITTIFSERFENDED GGSAVTSTGTIEDGEOPASCTGSEDSSEGAIS SESEENGSSAMDSTVAREGTNYPLVAAGPCD DEGIVTSTGAKEEDEEGEVVTSTGRGNEIGH ASTCTCLGEESEGVLICESAGEDSQGTVVBLOT GGSAVTSTGTTEDGEOPASCTGSEDSSEGAIS SESEENGSSAMDSTVAREGTNYPLVAAGPCD DEGIVTSTGAKEEDEEGEVVTSTGRGNEIGH ASTCTCLGEESEGVLICESAGEDSQGTVVBLOT DICSSAKGTVSESVTSAVSGKDEVTPYPGGCE GPMTSAASDOSDSQLEKVEDTTISTGLYGGS TDVLVSGEVPEGEVAHTSTSPEKEDEDUITSVE VEAEAGAAMMANENNYDSMSGTEKGKDT DICSSAKGTVSESVTSAVSGKDEVTPYPGGCE GPMTSAASDOSDSQLEKVEDTTISTGLYGGS TOVLVSGEVPEGEVEAHTSTSPEKEDEDUITSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIISTSTTINDYTPQVSAITDVEGGIISTSVEDC GFYSSAVPQEEGDPSVTPAERMGOTAMISTS TSGGCAVMIGGAVLQDEDATTITRVEDLSDA RSEEKDECAMISTSIGEEFELPSSATTIKCAES LQVVAAAVEERATGVVLISTADFEGPMPSAPP EASSPLASTSKERDECALISTSIAEECEASVS GVVVSESNERRAGTVMERENGGSIISTSSVEDC GGKESGVLANTSGEGIISTSSVEDC GGKESGVLANTSTEGERPSLPSAGGGIISTSSVEDC GGKESGCAVTICGAVATORALKADDMPPQGOTY GGKEPGPVLANTSEGGTAAGGSST THAGEENVALLANTGAIKADDMPPQGOTY GGKEPGPVLANTSEGGNANSPAHLRGPPGTTSGGTAK DSVSSIRYLAAAVNTGAIKADDMPPQGOTY	ì	ļ				•	,
ENKSDDKDGKEVDSSHEKARGNISMEKKL SRRLCENRRGSI SQBMAKGERLAANTI.STP SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE NVFEVSKTQDNINNISEQDIDSENKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAYSKMNPGE ENFIRIGTTEVNIDSETYEMRLISAPSENDRV QKNILKNTAAEHVAQGDATLEHSTNIDSSIPS LSSYTVYPLRESYPDPVIFIKRILSAPSENDRV QKNILKNTAAEHVAQGDATLEHSTNIDSSIPS LSSYTVYPLRESYPDPVIFIKRTLEGSTA STSPADHSALPNQSILTVRESEVILKTDSKERG GEFTVDTPAKASITSKRHIPEAHQATILDGKQ GKVIMPLGSKLTGVIVENENTKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKIDGKAVD HVVGLNTEKYAETVKLKKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQGTE  DVATORRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADGLIGT HSRNNPLHVGAEASECTVASAAEGGAVVTE GFAESETFILTSTKSGESGCAVAESSDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TERAGSISSEEVDGSQGNMMRMOFKKSTEG TVTCTGAGGRSDNFVICSTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDSEGFOVTSTGREBGHA SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDSEGEDVYTSTGRORBIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAANMAMENNVDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSKOEDVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTSCLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLISTSTTNDYTPQVSATITVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLARS RESEKDECAMISTSIGEBFISSATTIK CAES LQPVAAAVEERAGTPVLISTADFEGPMPSAP EAESPLASTSKEEREDGEALITSTVEAES LQPVAAAVEERAGTPVLISTADFEGPMPSAP EAESPLASTSKEEREDGEALITSTVEAES SEELDALSTSKEEREDGELITTREVELSDA AIISTSTAECMPISASDRHEENQLTAINEGECASVS GVVVESENEREGTVTREEFRADPSAVSGODSQLTANS REGCGEPSAAVSTEGENSTREEDSATTIK CAES LQPVAAAVEERAGTPVLISTADFEGPMPSAP EAESPLASTSKEERDGEALISTITICECLSA GGEFGPVLAVSTEEGENDROPSVHKPSAQGH PSAVCAEKEEKHGKECPEIGPFAGGGSST ASYSAGRGLEGNANSPAHLRGPECTSGTAK BUSSVSIKYLAAVNIGAKLADDMPPVQGTVA  GGKEPGPVLAAVSTEEGENDROPSVHKPSAQGH PSAVCAEKEEKHGKECPEIGPFAGGGST ASYSAGRGLEGNANSPAHLRGPECTSGGTAK BUSSVSIKYLAAVNIGAKADDMPPVQGTVA	J	}	٠.				
SRRLCENRRGSISQEMAKGEEKLAANTLSTP SGSSLQRPKKSGDMTLFEQEPMEIDSEPGVE NVFEVSKTQDNRINNSEQDIDSEMMKQKTS ATVQKDELRTCTADSKATAPAYKPGGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE KEPHRGTTEVMIDSETVHRMLLSAPENDRV QKNLKNTAAEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLGSTA STSPADHSALPPQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPRAHQATLLIGKQ GKVIMPLGSKLTGVEKENTLEGKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKNGGIAVD HVVGLNTEKYAETVKLKHKRSPGKNGGIAVD HVVGLNTEKYAETVKLKHKRSPGKNGGIAVD HVVGLNTEKYAETVKLKHKRSPGKNGGIAVD HVVGLNTEKYAETVKLKHKRSPGKNGOTE DVATOPRRABKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVQAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESDRAADL LAVHAVKLEANVNSVVTEKADDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDAVTSAG TEIRAGSISSEVDGSGOMMRMGPKKSTEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASGEGDSSVDGTF GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEARAGAAMNANENNYDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVGGGC GPMTSAASDQSDGLKKEDTTSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLISTSTTNDYTPQVSATDVEGGLSDALRTE ENMEGTRYTTEFFAPMPSAVSGIDSQLTAS RESEKDECAMISTISIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGRMPSAPP EAESPLASTSKEEKDECALISTIADFEGANSPAP EAESPLASTSKEEKDECALISTIADFEGANSPAP EAESPLASTSKEEKDECALISTIADFEGANSPAP EAESPLASTSKEEKDECALISTIADFEGANSPAP EAESPLASTSKEEKDECALISTIADFEGANSTALTSTAECKPSASTORPEGARSTIADFEGANSTALTSTAECKPSASTORPEGARSTIADFEGANSTALTSTAECKPSASTORPEGANSTIADFEGANSTALTSTAECKPSASTORPEGANSTARDFEGANSTALTSTAECKPSTATTSTAECKPSASTORPEGANSTALTSTAECKPSTATTSTAECKPSTATTSTAECCASVS GVVVESENERAGTVMEEKDGGGISTSSVEDC GGKFGPPLAASTGEGERMARADDMPPVQGTVA ANSTSTATCMARADDMPTAANSTALREGGTSTGTAK BUSHSTANTARADDARADDMPTAANSTA	i	Į			1		
SGSSLQRPKKSGDMTLIPEQEMEIDSEPGVE NVFEVSKTODNRNINSIDDENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAYSKMNGGE KEPRIRGITTEVNIDSETIVERMLLSAPSENDRV QKNILKNITAAEEHVAQGDATLEHSTNILDSSPS LSSVTVVPLRESYPDPVIFDKRTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKSPGKVKDISID VERRIENSEVDTSAAGSGAPSVLHQNOQTE DVATOPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADBGLIIGT HSRNNPLHVGAEASECTVFAAAEEGAVVTE GFAESTFILTSTKGEGSEGCAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRAGSEVGGTTITISEVESDAVTSAGSEE KCDGSLSRAGSEVGGTTITISEVESDAVTSAGSEE KCDGSLSRAGSEVGGTTITISEVSSDAVTSAG TEIRAGSISSEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAAGPEERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGEDVVTSTGRONEIGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVUGH ASTCTGLGEESGVLICESAAGDSQIGTVSG VDVLVSGEVECEVAHTSPSEKEDEDITISVE NEECDGLMATTASGDITNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAAPMSAVSGODSQLTAS RESECDECAAINTSISGEEFLPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGRMPSAPP EASSPLASTSKEEDDEGLISTSTAEECASPV GGVVESENERAGTVMEEKDGSGISTSSVEDC EGGYSSAVPQEEGDPSVTPAEEMGTAMINSTS TSEGCEAVMIGAVLQEDRITITIVEDLSDA AIISTSTAECMPISASIDRHEENQLTANIPEGN GGKEPGPVLAVSTEEGHNGPSVHRPSAGQGH PSAVCAEKEKHOKECPEIGPFAGGGKESTL HLINAEEKNVLLNSLOKEDKSPETGTAGGST ASYSAGRGLEGNANSFAHLREPGCJTGGTAK DSSVSSIKYLAAAVNTGAKADDMPPVQGTVA			Į.	ļ			
NYFEVŠKTODNRNNINSIQÜDSENIKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADIRSTLTKKMHIQSAVSKMNPGE KEPIHRGTTEVNIDSETVHEMLLSAPESDRIV QKNLKNTAAEBHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPLFPKRYTLGGSTA STSPADHSALPNOSLTVPESEVLKTSDSKEGG GGFTVDTPAKASTISKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDMA KKENDLNAEPNLKQTTKATVENGKKEGGLVDT HVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNOQTE DVATOPRRAEKTSVATSTEGKDKOVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFASESTFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDAVTSAG TERAGSISSEEVDGSGGNMRMGPKKSTEG TVTCTGAEGRSDNFVICSVTGAGFRERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSGIGTVVEH VEARAGAAIMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVGGCE GPMTSAASDQSBOLEKSLEDTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIJSTSTTNDYTPQVSATIVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGIDSQLTAS RESEKDECAMISTISIGEPISAGTIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP BAESPIA-STSKEKDED-LISTSTAEECGESAPP BAESPIA-STSKEKDED-LISTSTAEECGSAPS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTTAEEMGGTAMISTS TSGCCEAVMGGAVLGPERMPSAPP BAESPIA-STSKEKDED-LISTTAEECGESAPP GGKEFGPVLAVSTEEGHNOPSVHKPSAQGGH PSAVCAEKEEKHGKECPEIGPFAGGGAK AIISTSTAECMPISASIDRHEENQLTANIPEGN GGKEFGPVLAVSTEEGHNOPSVHRYSAQGGH PSAVCAEKEEKHGKECPEIGPFAGGGAK ASSAGRGLEGNANSFAHLREPGCJTSGGTAK DSSVSIKYLAAAVNIGAKADDMPPVQGTVA	1		ł	}	į		
ATVQKDELRICTADSKATAPAYKPGRGTGV NSNSEKHADHRSITLKMIIGSAVSKMIPGE KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEHVAQDATILEHSTNLDSSPS LSSVTVVVLRESVDPUPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG GEGFTVDTPAKASITSKHPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KENDLINAEPNLKQTIKATVENGKLDGLAVD HVVGLNTIEKYAETVSLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRAEKTSVATISGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSADSIEVGITTISISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNPVESVTGAGPREERMVT GAGVVLGDINDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAAPCD DEGIVTSTGAKEEDEEGEDVYTTGRGNEIGH ASTCTGLGESSGSVLICESAEGDSQIGTVVEH VEAEAGAAIMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDSVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NECCOGLMATTASGDITNQNSLAGGKNQGK VLIISTSTNDVTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSALTIKCAES LQPVAAAVEERATGPVLISTADEGRMPSAPP EAESPLAASTSKEKEKDECALISTISIAEECEASVS GVVVSGENERAGTVMEKENGOSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGSVLQDEDRLITITR VEDLEDA ALISTSTABCMPISSASIDRHEENQLTADNPEGM GGKEPGPVLAVSTEEGINGPSVHKPSAGQGH PSAVCAEKERHIGKETGERFEIPERSATTIKCAES LQPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGSVLQDEDRLITITR VEDLEDA ALISTSTABCMPISSASIDRHEENQLTADNPEGM GGKEPGPVLAVSTEEGINGPSVHKPSAGQGH PSAVCAEKERHIGHEGPITSGLAGNST ASYSAGRGLEGMANSPAHLRGPGTSGGTAK BSSVSSINTLAANVIIGAIKADDMPPVQGTVA	1		1				
NSINSEKHADHRSTLTKKMHIQSAVSKMNPGE KEPIHRGTTEVNIDSETVHRMLISAPSENDRV QKNILKNTAAEEHVAQQDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSGKVKDISID VERRNENSEVDTSAGGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAEEDRAADL LAVHAVKIEANNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMRMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVHDGTE GESAVTSTGITEDEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE EMMEGTRVTTEFFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVILSTADFEGMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPYSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA ALISTSTAECMPISSASTEREERDLTADNPEGG GCLSATERSKEEKDEPLOPPAGPGGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPGTSGQTAK BSSVSSTRYLAAVMIGAIKADDMPPVQGTVA	1		<b>j</b>	j	Į		
KEPHRGTTEVNIDSETVERMLISAPSENDRY QKNILKINTAAEEHYAQGATLEHSTILDSSPS LSSVTVVPLRESVDPDVPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG GEGFTVDTPAKASTISKRHPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNI.KGTVENGEKSGGAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGGSAFSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSADSSIVEGTITFIISEVESDGAVVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNPVCSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVANDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESSENGESAMDSTVAKGTIVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRONEIGH ASTCTGLGESEGGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKDEDIITSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIJSTSTTNDVTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSALTIKCAES LQPVAAAVEERATGPVLISTAGGKNQGK VLIJSTSTTNDVTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTIGGESFELPISSALTIKCAES LQPVAAAVEERATGPVLISTAGGKNQGK VLIJSTSTTNDVTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDSQLTASS RSEEKDECAMISTIGGESFELPISSALTIKCAES LQPVAAAVEERATGPVLISTADEGPMPSAPP EAESPLAASTSKEEKDECALISTSIAEECEASVS GVVVSENERAGTVMEEKDGSIGTSSVEDC EOPYSSAVPQEEOPSTPABEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASJDRHEENQLTADNPEGM GCLSATEVSKHKVPNELAINNNCCFGPVR GGKEPGPVLAVSTEEGINGPSVHKPSAQGGH PSAVCAEKEEHGISGPEIPPAAGROGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPGTTSGGTAK BSSVSSTRYLAANVITGGIKADDMPPVQGTVA	1	ŀ	[				
QKNILKNTAAEBHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRIIPEAHQATLLIDGKQ GKVIMPLGSKLTGVIVENENTIKEGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSVLPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESTFLTSTKEGGSGECAVESEDGAUTS KCDGSLSRDSEIVEGTITTISEVESDGAVTSAGS TEIRAGSISSEEVDGSGONMERMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGINDAPPGTSAGSGOSGNMMERMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGINDAPPGTSAGSGOGSSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFRAMPSAVSGDDSQLTAS RESEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP BAESPLASTSKEEKDECALISTSIAECCEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC GGVVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDLTITTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHOPSVIHRPSAQGGROSESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLARPEQTSGOTAK	1	1				ļ	NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE
SSYTVVPLRESYDPDVIPLEDKTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEINTKEGGLVDMA KKENDLNAEPHLKGTIKATVENGKENGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINENSEVDTSAGSOSAFVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSNDSEIVEGTITFISEVESGDAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEBEDEGEDVVTSTGRGNEIGH ASTCTGLGEESGVLICESAEGDSQIGTVVEH VAEAGAAINMANENNVDSMGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITSTGLVGGS YDVLVSGGVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTNQNSLAGGKNQGK VLIISTSTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAVEERATGPVLISTADEGPMPSAPP EASSPLASTSKEEKDECALISTSLAEECEASVS GVVVESENERAGTVMEEKGGGIISTSSVEDC EGFVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTIT VEDLSDA ALISTSTACMPISASDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGSVIKKPSAQGGH PSAVCAEKEEKHÖECEGFOFPAGGOKESTL HLINABEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLREPRQTSGOTAK	1	ł			ľ		KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV
STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHPEAHQATILLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPILKQTIKATVENGKKDGJAVD HVVGLNTEKYAETVKHKKRSFGKVKDISID VERRNENSEVDTSAGSGSABSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRGSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAEEGGAVVTE GFAESETPLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTIEVESDGAVTSAG TEIRAGSISSEEVDGSGGNMMEMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPGTSAQEGGESVNDGTE GESAVTSTGITEDGEOPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEGEDVVTSTGREGEIGH ASTCTCLGEESEGVLICESAEGDSQGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGFVESSVTSAVSGKDEVTPVPGGCE GFMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAFMPSAVSCDDSQLLTAE  SESELDECAAMISTSIGEEFELPISSATTIKCAES LQPVAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC GGVSSAVQFGEGDPYTTAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTITR VEDLSDA AIISTSTAECMPISASIDHENENQLTADNPEGN GOLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEGGINGFSVHKPSAGQGH PSAVCAEKEEKHGEGGIISTSSVEDC GGKEPGPVLAVSTEGGINGFSVHKPSAGQGH PSAVCAEKEEKHGERGISTSTSVEDC GGKEPGPVLAVSTEGGINGFSVHKPSAGQGH PSAVCAEKEEKHGECEIGPPAGROGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRDPEQTSGOTAK	1		ļ	Į	<u> </u>		QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
EGFTVDTPAKASITSKRHIPEAHQATILLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVHQRNGGID VERRIENSEVDTSAGSGSAPSVHQRNGGID DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITISSETRQSEVALPCTSIEADBGLIGT HSRNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVGTITTISEVESDGAVTSAG TEIRAGSISSEVDGSQGMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMY GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSBEKDBCAMSTSIGGEFELPISSATIKCAES LQPVAAAVEERATGFVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITITRVEDLSDA AIISTSTAECHERSLDTHEENGLTANDFEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVTIKPSAGQGH PSAVCAEKEEKHGKECPEIGFPAGRGOKESTI. HLINAEEKNYLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGQTAK DSSVSIRVLAAAVILGALKADDMPPVQGTVA			į	1	ł		LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
EGFTVDTPAKASITSKRHIPEAHQATILLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVHQRNGGID VERRIENSEVDTSAGSGSAPSVHQRNGGID DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITISSETRQSEVALPCTSIEADBGLIGT HSRNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVGTITTISEVESDGAVTSAG TEIRAGSISSEVDGSQGMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMY GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSBEKDBCAMSTSIGGEFELPISSATIKCAES LQPVAAAVEERATGFVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITITRVEDLSDA AIISTSTAECHERSLDTHEENGLTANDFEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVTIKPSAGQGH PSAVCAEKEEKHGKECPEIGFPAGRGOKESTI. HLINAEEKNYLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGQTAK DSSVSIRVLAAAVILGALKADDMPPVQGTVA				[	ļ		STSPADHSALPNQSLTVRESEVLKTSDSKEGG
GKVIMPLOSKLTGVIVENENTKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITISSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESEFFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGMMMRGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSSAGEGDGSVDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESNERRAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDERALITITRVEDLSDA AIISTSTAECHPISASTIRGERGEASST GGKEPGPVLAVSTEEGHNOPSVHKPSAGQGH GGKEPGPVLAVSTEEGHNOPSVHKPSAGQGTA GGKEPGPVLAVSTEEGHNOPSVHKPSAGGDST HLNABEKNYLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTTSGGTAK DSSVSSIRVLAAVMIGAIKADDMPPVQGTVA	· ·			i	Ì		
KKENDLNAEPNILKQTIKATVENOKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHORNGQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSATIDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSLAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHGKECPEIGPFAGRQKESTI HLINABEKRVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTTSAGCST	1			(	1		,
HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSASQEGDGSVDDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGEGDVVTSTGRGPEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQK VLIISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAHPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESSENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPPVR GGKEPGPVLAVSTEEGHNOPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGGKESTI. HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	l			1		
VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVQAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSKEDEDJITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEFFEAPMPSAVSGDDSOLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTFAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTI. HLINABEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNIGAIKADDMPPVQGTVA		ļ		]		}	
DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRGSIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGFASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTDYTPQVSAITDVEGGLSAALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMGAVLQDEDRITITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHRPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTIL HLINAEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRVLAAVMIGAIKADDMPPVQGTVA	1	ł	1		Į		,
AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAASECTVFAAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVVEGTITTISEVESDGAVTSAG TEIRAGSISSEVDGSVGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGGE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRYTTEEFFAAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDERLITTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGECPEIGPFAGRGQKESTL HLINABEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRVLAAVMIGAIKADDMPPVQGTVA	1		ļ		1		
HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEGRAADL LAVHAVKEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQOMMMRMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVAQEEGDPSVTPAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASDDRIEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHRPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTI. HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK	1	l	}	}	J	ļ	_
GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GFMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEBGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA						i	l
LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITIFISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMMRMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVL.GDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEENNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSIRYLAAVNTGAIKADDMPPVQGTVA	1	ļ	ì				4
KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RESEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEENVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSIRYLAAVNTGAIKADDMPPVQGTVA	ł	l	]	l	ł	ľ	!
TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAGAGRSDNFVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RESEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		l	1		ļ		l
TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKCPEIGPFAGRGQKESTL HLINAEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	}	}			}	]	,
GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGROEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDJITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ì					
GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINABEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1			ļ		
SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGROREIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALATE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		1	1	1	i	· · · · · · · · · · · · · · · · · · ·
DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSSTL ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1		[		1	GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS
ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1	Į	1	!	}	SESEENGESAMDSTVAKEGTNVPLVAAGPCD
VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTIL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1			1		DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH
VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTIL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		.]	1			ASTCTGLGEESEGVLICESAEGDSQIGTVVEH
DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1	(	1	1	(	,
GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1		1			
YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1	}	l	1	}	
NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1	}	İ	1	l	
VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					1	ĺ	
ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			1		1		
RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTIL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	1	1	1		
LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITITR VEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHIGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTIL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		1	1	1		
EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		l	1	1	1	
GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	1	1	1	1	
EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITR VEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	}	j.	}	1	1	
TSEGCEAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTI HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		l		1	l .	-
AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	i	1	1	1 .	
GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	(	1	1	1	1	1	
GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		1	1		1	
PSAVCAEKEEKHGKECPEIGPFAGRGQKESTIL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		1	}	1	1	GDLSATEVSKHKVPMPSLIAENNCRCPGPVR
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		]				1	
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		l	1		1	1	PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL
ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1	İ	<b>{</b>	I	{	
DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	(	1	1	Ī	ĺ	[	7
		1	1	1	1	1	
1 Selection of the sele	1	}	l		1	1	
	L	ــــــــــــــــــــــــــــــــــــــ		L	<del></del>	<del> </del>	

SEQ ID	SEQ ID	Mct	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	,	<u> </u>		peptide		/=possible nucleotide deletion, \=possible
'				sequence		nucleotide insertion
					1	SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK
			l l		1	WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS
	i	1	į	Į		SEENVCDIGNEESPLNVLGGLKLKANLKMEA
			ĺ	Į	İ	YVPSEEEKNGEILAPPESLCGGKPSGIAELQRE
	1	ł	1			PLLVNESLNVENSGFRTNEEIHSESYNKGEISS
		1	ì	ļ	İ	GRKDNAEAISGHSVEADPKEVEEEERHMPKR KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC
	]	ļ	]	ì	1	PETEPHATKEENSRDLEELPKTSSETNSTTSRV
		İ			Į.	MEEKDEYSSSETTGEKPEQNDDDTIKSQE
		<u> </u>	1016		842	EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH
1302	2652	A	10167	321	842	FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG
	Į.					NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE
	1	1			1	DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP
	ſ	1		}	1	EARETKCYVRSSVGCVEPLTTQAEVTENLDR
		1	1	ļ		KNSQQVFKLLKKK
1303	2653	A	10171	206	429	NMILLKKRRLLINSLGEGTINGLLDELLETNV
1303	2033	1	10171	200	127	LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA
		1				RACEICITYI
1304	2654	A	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF
1304	2054	1"	10101	1 - 10		YHEAVVLFTQALKLNPQDHRLFGNRSFCHER
		1	1		1	LGQPAWALADAQVALTLRPGWPRGLFRLGK
	)	}	1	1	ļ	ALMGLQRFREAAAVFQETLRGGSQPDAAREL
			1	1	1	RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA
		ŀ				ELAPSGLPSLRCPRSTALRSPGLSPLLH
1305	2655	A	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ
			1			SDFLTPPVGGAPWAVATTVVMYPPPPPPPHR
}	1	1	İ		1	DFISVTLSFGESYDNSKSWRRRSCWRKWKQL
Ì		1	Ì			SRLQRNMILFLLAFLLFCGLLFYINLADHWKG
ļ		1	-			IRNTCT
1306	2656	A	10195	1	410	IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY
		l	İ			NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI
}			ì	1		GIDDEDDSTFTITVDQKTFHFQARDADEREK
		1		ŀ		WIHALEETILRHTLQLQVRVFTWFPDSSLVGA
			<u> </u>	<u> </u>		FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM
	İ					DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS
			1	<u> </u>	<u> </u>	SPDQQNYTKSR
1308	2658	A	10214	2	453	ECGGIRQPGPPPPALASAPAATMNRVGGSPS
1		-	1	1		AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP
		1	1			DNITQVMSLHTQYLESFLRSQFYMLRMDGPL
	1	1	1	1		PLPYRHYIAIMAAARHQCSYLINM
	<u> </u>	1	1.000	1,	401	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP
1309	2659	A	10233	45	421	RALESRTFQGSERSRWGPPLESTKENVQCGH
}	1	1	i	Į.		RPAFPNSSWLPFHERLQVQNGECPWQVSIQM
1		1	1			SRKHLCGGSILHWWWVLTAAHCFRRTLLDM
				ł	ļ	AV
		4	10041	1242	442	AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK
1310	2660	A	10241	243	442	HKKGOSAEIOKKRTRRAFKFQRAITGASLADI
		1	(		1	MAK
1000	1300	+	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ
1311	2661	Α	10261	131	170	VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ
ì		1	1	ì		LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF
1		1	1	İ		OKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI
1	}		-	ľ		DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML
1			1	}		OMVCAOCHGNFCIQHRHPLDHSCRHGSRPTI
[			1	1	1	KAG
1	1					
1312	2662	A	10270	1 3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP

						(1-A)-i (2-C
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ì	Į.	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	]		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ì	l	ļ	peptide		/=possible nucleotide deletion, \=possible
	1	i	ł	sequence		nucleotide insertion
					<del></del>	LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF
	1					EHVCKPKVEHLIGYEVFECTHNMAYMLKKL
	1		1			LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV
	Į	ł	1		1	REESSFSDIPDVKNDFAFLLHMVDQYDQLYS
	1	i	1	l		KRFGVFLSEVSENKLREISLNHEWTFEKL
		<del> </del>	10000	1001	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR
1313	2663	A	10287	1221	200	VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL
	1	1		ì	1	
	1	1	Į.			LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE
	Ì		1	1	[	EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD
	1					GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL
			1	İ		KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL
	i	1	1	1	1	PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW
,	1	1	1	-8-	1	GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF
1	0.0			}	}	PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL
		1			1	LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL
1314	2004	1.7	10200		1.054	VKKDKSEKELKALCIDQLDVFLQKETQIFVEK
i	1	1	1		i	LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK
	1	1	1		1	DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR
į	1	1	1	1	· ·	ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD
	1			1	1	RYNRRGRSRSYSRSRSRSWSKERLRERDRD
}	1	i			}	RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN
	1	Į.		1		RSKI KSKSK I KSKERDL VAPA I DEDATOPLEN
1	1	1	i		1	NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG
		ĺ	1		ļ	NNTTESWSEFHEDQVDHNSYVRPPMPKKRC
	İ	1		1		RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN
l	1.	1	1	į		LPGMQPFPAQPPVVEGPPPPGLPPPPPILTPPPV
1	ì	Į.				NLRPPVPPPGPLPPSLPPVTGPPPPLPPLQPSG
	l l			ĺ		MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT
	ł		į	į	1	ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR
i	1	i				AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAEERRTRGA
1313	2003	A	10293	1 447	, 1551	GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI
1			l l	1		GTPSPCGSAQAAAAAAEEATEKIPALRPALL
					}	WALLALWLCCATPAHALQCRDGYEPCVNEG
1		1		1	į	MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE
	1	-		1		KNRCQNGGTCVAQAMLGKATCRCASGFTGE
1	1	1		1		DOONGTOIDGEVERDOT NICOTOLINI EDUTVE
		1				DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE
		-	1	1	1	CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS
1	1	1		l .		GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT
	1					LWVMVSGATSTSTSTL
1316	2666	A	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG
		1				YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI
1		1				GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF
	1				[	NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR
1	1	1		1	1	NGRLLGVCASVDNCRLFVGGIPKTKK
1215	2007		10201	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ
1317	2667	Α	10301	156	1550	DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP
			1	1	1	LPAASSGMKSSKSSTSLAFESRLSRLKRASSE
1		1	1	1	1	DTLNKPGSTAASGVVRLKKTATAGAISELTES
				1		
1		1	1	1	1	RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV
	1	- 1			}	PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP
1	1	1			ļ	KQENEGGEKAALESQVRELLAEAKAKDSEIN
1	1	1			1	RLRSELKKYKEKRTLNAEGTDALGPNVDGTS
	į	1	1			VSPGDTEPMIRALEEKNKNFQKELSDLEEENR
1		1	1		1	VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ
1	1	1	1	ł	ì	ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG
	1	- 1	1	1	l l	ACCOUNTS AND ACCOUNTS A DESCRIPTION OF THE COMME
1		J	}	l l	i i	SSSSDVTKASLSPDASDFEHITAETPSKPLSSTS
		-			į	SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS
						SSSSDVTKASLSPDASDFEHTTAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

PCT/US01/03800

SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of nucl-	NO: of peptide	hod	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uonoo	(	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
[		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Į.	1	ł	1	peptide	1 -	/=possible nucleotide deletion, \=possible
	ļ	Į.		sequence		nucleotide insertion
	<u> </u>	<del>                                     </del>				NEKLVDEKTILETSFHQHRERAEQLSQENEKL
İ		1	1		Į.	MNLLQERVKNEEPTTQEGKHELEQKCTGILE
	'	Ì	1			QGRFEREKLLNIQQQLTCSLRKVEEENQGAL
ł	•	]				EMIKRLKEENEKLNEFLELERHNNNMMAKTL
İ	į.		1			EECRVTLEGLKMENGSLKSHLQG
1318	2668	A	10303	333	879	GECFIMAAVVQQNDLVFEFASNVMEDERQL
		(		ĺ	[	GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
1	1	ł	1	1		NDMITESSLDVAEEEIIDDDDDDTTLTVEASCH
1		1	ł	ł	1	DGDETIETIEAABALLNMDSPGPMLDEKRINN
						NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ
		}				QVQEKYADSPGASSPEQPKRKKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLIIATALSL
	1	1		1	1	LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR
1	1	1	1	1	1	LPRTLAGLLAGGALGLAGALMQTLTRNPLAD
1		[	1	1		PGLLGVNAGASFAIVLGAALFGYSSAQEQLA
1		i	1	ŀ	1	MAFAGALVASLIVAFTGSQGGGQLSPVRLTL
}	1	1			<u> </u>	AGVXL
1320	2670	Α	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV
		1	1:	1	ļ	AVVDIQSDKAANVAQEINAEYGESMAYGFG
1		1		ļ.	1	ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI
			1	1		AKAAFISDFQLGDFDRSLQVNLVGYFLCARE
[	-	1 _	<u> </u>	<u> </u>		FSRLMIRDGIQGRIIQINSKSDE
1321	2671	Α	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY
			į	1	ĺ	AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD
1 '		İ	ſ		1	ILKCTLLVFGVRILYILKLNYTTEECDMKNMH
	Ì					YVDPDHVKRAQKYAQQVLQKESPPKFAKTS
L					<u> </u>	MALLFEHRYSVDLLPFVQKAPTDSEA EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS
1322	2672	Α	10333	25	423	ERKMRAHQVLTFLLLFVITSGASENASTSRGC
		1	j		1	GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA
						LITLLIMLILLGRLPFIKEKEKKSPAVLHFLFL
1		ļ		ļ	]	LGTLG
		<del> </del>	10504	<del></del>	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH
1323	2673	A	10334	52	420	QVLTFLLLFVITSVASENASTSRGCGLDLLPQ
1	-	1	ľ		1	YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI
	ļ	1	1			LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
100	0.51	+	10226	<del>  ,                                   </del>	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE
1324	2674	A	10336	1	932	NSVTHHEVKCQGKPLAGIYRKREEKRNAGN
i		,				AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA
				1	1	AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA
1		-		1	1	PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE
			1	1	1	LQSEERKRIDELIESGKEEGMKIDLIDGKGRG
1		1	1	j		VIATKOFSRGDFVVEYHGDLIEITDAKKREAL
1						YAQDPSTGCYMYYFQYLSKTYCVDATRETN
-		1		1	1	RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS
			Ì	1		RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	+	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL
1323	2013	1^	1,0556	1		RGVTATFGRPAEWPGYLSHLCGRSAAMDLG
1		- [		1		PMRKSYRGDREAFEETHLTSLDPVKQFAAWF
	1	1		1		EEAVQCPDIGEANAMCLATCTRDGKPSARML
{		1		1		LLKGFGKDGFRFFTNFESRKGKELDSNPFASL
		1		1	1	VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS
1		1		1		RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE
- {		- (	1	1		OLYODOEVPKPKSWGGYVLYPQVMEFWQG
ļ	ŀ			1		OTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE
1	1			1		DWLYERLAP
1326	2676	A	10344	12	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV
1340	20/0	^	10344	1-	1 -5.	LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT
	- 1	1	1	1		LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA
1		- 1	1	· L		

CDO ID	CT 022	Man	CEO	Dendistad	Predicted end	Amino acid sequence (A-Alanine C-Cysteine,
SEQ ID	SEQ ID	Met hod	SEQ ID NO:	Predicted beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	nou		nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ĺ	in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	l	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Ì	1	residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
			ļ	1	sequence	/=possible nucleotide deletion, \=possible
		}	j	peptide	}	, .
				sequence		nucleotide insertion
		1				HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD EDOGYVDLVIKVYLKGVHPKFPEGGKMSQY
'		1				
	Į.	1	1	ŀ	Ì	LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP
i	{			İ	1	NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA
	1	1		1		ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ
	ł	ł	ł	Ì	1	ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD
ł	ł	ł			l .	MIREHLPAPGDDVLVLLCGPPPMVQLACHPN
		<u> </u>		<u></u>		LDKLGYSQKMRFTY
1327	2677	Α	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR
1		l l	1	1	1	RYHRLSDMSMLAERRRKQKWAVDPQNTAW
1	1	l		<u>}</u>	j	SNDDSKFGQRMLEKMGWSKGKGLGAQEQG
i		i	ì	l	i	ATDHIKVQVKNNHLGLGATINNEDNWIAHQ
	}	ļ	l	,		DDFNQLLAELNTCHGQETTDSSDKKEKKSFS
{	ľ	l	1		1	LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL
				Ì	į	DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF
}	1	1	}		1	TIQEYFAKRMAALKNKPQVPVPGSDISETQVE
	İ		1	•	İ	RKRGKKRNKEATGKDVESYLQPKAKRHTEG
1	1	}				KPERAEAQERVAKKKSAPAEEQLRGPCWDQ
		i			_	SSKASAQDAGDHVQPA
1328	2678	A	10346	173	439	GSAAMKVKIKCWNOVATWLWVANDENCGI
1	1	1		ì		CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF
ļ	1	l	İ	1	1	HMHCILKWLHAQQVQQHCPMCRQEWKFKE
1329	2679	A	10351	3	964	OMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL
1.5-5	-0.,					SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN
1		1	1		Į	LSFADICVTSTTIPKMLMNIQTQNKVITYIACL
				1		MOMYFFILFAGFENFLLSVMAYDRFVAICHP
		1		1	ŧ	LHYMVIMNPHLCGLLVLASWTMSALYSLLQI
Ì	İ	1		1	1	LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF
-	İ		1		1	LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH
1		1 .			1	AISSAQGKYKAFSTCASHLSVVSLFYGAILGV
	1	1		1	i	YLSSAATRNSHSSATASVMYTVVTPMLNPFI
		1			Į	YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	1 A	10352	34	2573	IPFLKSCCCCLFDFPPPPLDQVQEEECEVERV
1330	2080	^	10332	] 34	] 23,3	TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE
1	1			1		TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL
1			1		1	FYTERAHVRTLKVLDQVFYQRVSREGILSPSE
		1	1	Í	1	LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS
Ì	ì	1	,		1	VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ
		ļ	1	1	1 -	PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR
		1	1	1	1	LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT
1		1		1	1	EREKVKKAADHCROILNYVNOAVKEAENKO
		1	İ			RLEDYORRLDTSSLKLSEYPNVEELRNLDLTK
		1		1	1	
1	1	1	1		í	RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV
1		1				LLQKQDDRLVLRCHSKILASTADSKHTFSPVI
1	1	1	1			KLSTVLVRQVATDNKALFVISMSDNGAQIYE
1	1	1				LVAQTVSEKTVWQDLICRMAASVKEQSTKPI
1	İ	1		}		PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL
]	1	1		1	1	QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD
1	1	1				LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS
1	1	1		1	1	HLPVSEERWALDALRNLGLLKQLLVQQLGLT
1	1	1.	1			EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE
	1	1		1	i	NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE
1		1	1	1	1	SFAPRDSVGLAPQDSQASNILVMDHMIMTPE
{		1	1	1		MPTMEPEGGLDDSGEHFFDAREAHSDENPSE
j		1		1	1	GDGAVNKEEKDVNLRISGNYLILDGYDPVQE
		1	1	1	ļ	SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP
}	1		1	1	1	QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI
	{	1			1	OSPSSCADSQSQIMEYIHKIEADLEHLKKVEE
1		1	i	1	Į.	SYTILCORLAGSALTDKHSDKS
i			_1		<u> </u>	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end on nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1331	2681	A	10353		2100	AVEFAEGALTMAPWPELGDAQPNPDKYLEG AAGQQPTAPDKSKETNKTDNTEAPVTKIELLP SYSTATLIDEPTEVDDPWNLPTLQDSGIKWSE RDTKGKILCFFQGIGRLILLLGFLYFFVCSLDIL SSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIG VLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP IIMGANIGTSITNTIVALMQVGDRSEFRRAFA GATVHDFFNWLSVLVLLPVEVATHYLEIITQL IVESFHFKNGEDAPDLLKVITKPFTKLIVQLDK KVISQIAMNDEKAKNKSLVKIWCKTFTNKTQ INVTVPSTANCTSPSLCWTDGIQNWTMKNVT YKENIAKCQHIFVNFHLPDLAVGTILLILSLLV LCGCLIMIVKILGSVLKGQVATVIKKTINTDFP FPFAWLTGYLAILVGAGMTFIVQSSSVFTSAL TPLIGIGVITIERAYPLTLGSNIGTTTTAILAAL ASPGNALRSSLQIALCHFFFNISGILLWYPIPFT RLPIRMAKGLGNISAKYRWFAVFYLIFFFLIP LTVFGLSLAGWRVLVGVGVPVVFIIILVLCLR LLQSRCPRVLPKKLQNWNFLPWMRSLKPW DAVVSKFTGCFQMRCCCCCRVCCRACCLLC GCPKCCRCSKCCEDLEEAQEGQDVPVKAPET FDNITISREAQGEVPASDSKTECTAL
1332	2682	A	10354	30	1377	SQQGSQPHRQGPPSLLTAPHSLDLPALPPGPR GSQGKLRRVLVPMSVKPSWGPGPSEGVTAVP TSDLGEIHNWTELLDLFNHTLSECHVELSQST KRVVLFALYLAMFVGLVENILLVICVNWRG SGRAGLMNLYILNMAIADLGIVLSLPVWMLE VTLDYTWLWGSFSCRFTHYFYFVNMYSSIFF LVCLSVDRYVTLTSASPSWQRYQHRVRRAM CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLFM APFETYSTWALAVALSTTILGFLLPFPLITVFN VLTACRLRQPGQPKSRRHCLLLCAYVAVFV MCWLPYHVTLLLLTLHGTHISLHCHLVHLLY FFYDVIDCFSMLHCVINPILYNFLSPHFRGRLL NAVVHYLPKDQTKAGTCASSSSCSTQHSIIIT KGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP TQPLTPS
1333	2683	A	10358	2	884	AAGAGADGREPASERASRAEPPAVAMGQND LMGTAEDFADQFLRVTKQYLPHVARLCLIST FLEDGIRMWFQWSEQRDYIDTTWNCGYLLA SSFVFLNLLGQLTGCVLVLSRNFVQYACFGLF GIIALQTIAYSILWDLKFLMRNLALGGGLLLL LAESRSEGKSMFAGVPTMRESSPKQYMQLGG RVLLVLMFMTLLHFDASFFSIVQNIVGTALMI LVAIGFKTKLAALTLVVWLFAINVYFNAFWT IPVYKPMHDFLKYDFFQTMSVIGGLLLVVAL GPGGVSMDEKKKEW
1334	2684	A	10367	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPQLLFP ELPLPHVPGQESAKRRSARRFLIMSELTKELM ELVWGTKSSPGLSDTIFCRWTQGFVFSESEGS ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSS WRDCSQEEQKELLCHTLCDILESACCDHSGS YCLVSWLRGKTTEETASISGSPAESSCQVEHS SALAVEELGFERFHALIQKRSFRSLPELKDAV LDQYSMWGNKFGVLLFLYSVLLTKGIENIKN EIEDASEPLIDPVYGHGSQSLINLLLTGHAVSN VWDGDRECSGMKLLGIHEQAAVGFLTLMEA LRYCKVGSYLKISKIPYLDCLASETHLTVFFA KDMALVAPEAPSEQARRVFQTYDPEDNGFIP DSLLEDVMKALDLVSDPEYINLMKNKLDPEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LGIILLGPFLQEFFPDQGSSGPESFTVYHYNGL KQSNYNEKVMYVEGTAVVMGFEDPMLQTD
1335	2685	A	10375	82	2929	DTPIKRCLQTKWPYIELLWITDRSPSLN  TRTKRRLGREKAMASPPRGWGCGELLPFML LGTLCEPGSGQIRYSMPEBLDKGSFVGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDINDNFPRFRDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISETFQLDSNLGEISTL QSLDYEESRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGLFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFIVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAAVSCVFLAFVIVLLV LRLRRWHKSRLLQAEGSRLAGVPASHFVGV DGYRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSEESCEKSEPLLMSDKVDANKEERRV QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT
1336	2686	A	10379	1	557	WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNAAGKRDGKAPAGGNGN KKSGKKEKK RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEK
						LAKLQAQVRIGGKGTARRKKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	A	10380		1263	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITEECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QPN
1339	2689	A	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690	A	10388	113	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFRIRCSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGL IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP ND
1341	2691	A	10392		5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEBELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTTVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion  SQVGGKRFECKDCGETFNKSAALAEHRKIHA RGYLVECKNQECEEAFMPSPTFSELQKIYGK DKFYECRVCKETFLHSSALIEHQKIHFGDDKD NEREHERERERERGETFRPSPALNEFQKMYG KEKMYECKVCGETFLHSSSLKEHQKIHTRGN PFENKGKVCEETFIPGQSLKRRQKTYNKEKLC DFTDGRDAPMQSSELSEHQKIHSRKNLFEGR GYEKSVIHSGPFTESQKSHTTTRPLESDEDEKA FTISSNPYENQKIPTKENVYEAKSYERSVIHSL ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS HNGNELVESNEKGESSIYISDLNDKRQKIPAR ENPCEGGSKNRNYEDSVIQSVFRAKPQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAKKK YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ ECGECFAHSSDLTEHQKIHDREKPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DPRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSSILPFHQRIHEQDQLYSM KGCDDGFIALLPMKPRNRAAERNPALAGSA IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
			,			QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ
1342	2692	A	10393	2	1350	LFNDHLSLARHQNTHTG GRPRSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQL LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE
1343	2693	A	10394	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC
1344	2694	A	10395	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR

SEQ NO:	of NO: of	hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotid		´	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uenc		l l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	·	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	- 1			peptide		/=possible nucleotide deletion, \=possible
.	. 1	1		sequence		nucleotide insertion
						RKLQGKLPELQGVETELCYNVNWTAEALPSA
1	ľ	1	·	ł	ĺ	EETKKLMWLFGCPLLLDDVARESWLLPGSN
ŀ	Ì	-	İ	1		DLLLEVGPRLNFSTPTSTNIVSVCRATGLGPV
1	[	1				DRVETTRRYRLSFAHPPSAEVEAIALATLHDR
1		i		ļ		MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR
ĺ		ĺ		Ì	į	LALEKANQELGLALDSWDLDFYTKRFQELQR
			ļ			NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDG
	1					QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA
1	1	[				IQGKEVRFLRPEDPTRPSRFQQQQGLRHVVFT
1				ļ		AETHNFPTGVCPFSGATTGTGGRIRDVQCTG   RGAHVVAGTAGYCFGNLHIPGYNLPWEDLSF
i	,		}			OYPGNFARPLEVAIEASNGASDYGNKFGEPV
	Ī					LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS
1				1	,	MEADHISKEAPEPGMEVVKVGGPVYRIGVGG
1	1	İ	1			GAASSYQVQGDNTSDLDFGAVQRGDPEMEQ
	Ì			1	J	KMNRVIRACVEAPKGNPICSLHDQGAGGNG
1	- 1					NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW
}	ļ	1				GAEYQESNALLLRSPNRDFLTHVSARERCPA
	İ					CFVGTTTGDRRIVLVDDRECPVRRNGQGDAP
1	1				•	PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP
	ł	İ			}	MLQPLALPPGLSVHQALERVLRLPAVASKRY
	1	- (				LTNKVDRSVGGLVAQQQCVGPLQTPLADVA
}.	j			ļ		VVALSHEELIGAATALGEQPVKSLLDPKVAA
1	l l	1				RLAVAEALTNLVFALVTDLRDVKCSGNWM
1		-	1	l	ļ	WAAKLPGEGAALADACEAMVAVMAALGVA
1	1					VDGGKDSLSMAARVGTETVRAPGSLVISAYA
1	ĺ				1	VCPDITATVTPDLKHPEGRGHLLYVALSPGQ
1	}	j	}		İ	HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM
	ļ	1		1		AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV
1	l	1	1	1	1	LEVQEPDLAQVLKRYRDAGLHCLELGHTGE
			1			AGPHAMVRVSVNGAVVLEEPVGELRALWEE
İ	1	Ì		1		TSFQLDRLQAEPRCVAEEERGLRERMGPSYC
1	1			]	]	LPPTFPKASVPREPGGPSPRVAILREEGSNGDR
1	i			Ì	•	EMADAFHLAGFEVWDVTMQDLCSGAIGLDT
ļ	1		1	1	}	FRGVAFVGGFSYADVLGSAKGWAAAVTFHP
1	İ			1	· ·	RAGAELRRFRKRPDTFSLGVCNGCQLLALLG
			1			WVGGDPNEDAAEMGPDSQPARPGLLLRHNL
				}	1.	SGRYESRWASVRVGPGPALMLRGMEGAVLP
				1		VWSAHGEGYVAFSSPELQAQIEARGLAPLHW
			}	ł	1	ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR
			1		1	HLAVMPHPERAVRPWQWAWRPPPFDTLTTS
13.	- 2002		10000	65	642	PWLQLFINARNWTLEGSC
134	5 2695	Α	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFQHRE
	ļ	1			1	RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI
		)	1	1	}	LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI
	٠	1	1		1	YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI
1		- {	1	1	1	MYLVLVLAVQVHAWQLYYSKKLLDSWFTST
				1		QEKKHK
134	6 2696	- A	10398	1	718	DDFVRCGPQSAAMGASARLLRAVIMGAPGS
134	2070	1.	1.0370	1.	1	GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT
					ļ	EIGVLAKAFIDOGKLIPDDVMTRLALHELKNL
		1		}	*	TOYSWLLDGFPRTLPQAEALDRAYQIDTVINL
1			1			NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK
				i	Į.	TVGIDDLTGEPLIQREDDKPETVIKRLKAYED
]		Į		1	1	QTKPVLEYYQKKGVLETFSGTETNKIWPYVY
L				<u> </u>		AFLQTKVPQRSQKASVTP
134	7 2697	A	10402	153	1969	KHRQENNALDMAPEIHMTGPMCLIENTNGEL
1		- 1	1	1		VANPEALKILSAITQPVVVVAIVGLYRTGKSY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
001100		l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
[	{	1	'	peptide		/=possible nucleotide deletion, \=possible
	1	ì	<b>\</b>	sequence	!	nucleotide insertion
<del> </del>	<del> </del>	<del></del>	<del>                                     </del>	- Saquessa	<del></del>	LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
}		1	1		ļ	PHPKKPEHTLVLLDTEGLGDVKKGDNQNDS
Ì	ł	ì	1	1	,	WIFTLAVLLSSTLVYNSMGTTNQQAMDQLYY
	1	1		1	ļ	VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
}	1	l	1		Į.	WTLRDFSLDLEADGQPLTPDEYLEYSLKLTQ
ŀ	1	1				GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
1	1	1	}	į	1	HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
	İ	1				FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
1	}	1	]	}		GDLPCMENAVLALAQIENSAAVQKAIAHYD
1	1	1	I			QQMGQKVQLPAETLQELLDLHRVSEREATEV
	İ	1	1	1		YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
1	1	1	1			ONOEASSDRCSALLQVIFSPLEEEVKAGIYSK
			1	ļ		PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
1 -	1	1.	i	1	1	OTYLKSKESVTDAILQTDQILTEKEKEIEVEC
ļ	1		j.			VKAESAOASAKMVEEMQIKYQQMMEEKEKS
	1.	1	1	ĺ		YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
}	1	1	1	1	į.	QEQARVLKERCQGESTQLQNEIQKLQKTLKK
1			1	1	1	KTKRYMSHKLKI
	1000	<del> </del>	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
1348	2698	A	10404	3	072	VAGGAPRRTTPVTMWRLLARASAPLLRVPLS
	1 .	]	j		}	DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
1	· ·	1	1			RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
	1			ì	1	EGNFAILALGGGYLHWGHFEMMRLTINRSM
	1	ĺ	1	1	į	DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1		1	-	1	i	GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
1	1	1		1		FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
i	i i		1		1	
		•			n. L.	DEDNING NIDWITEED LATANIMI GIRKVI SPVDI
			1	1 6	1	RERNNQNPWTFERIATANMLGIRKVLSPYDL
	0600	<u> </u>	10400		1104	THKGKYWGKFYMPKRV
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK  AGKLQSQLRTTVVAAAAFLDAFQKVADMAT  NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS  ALIDCLINPLQEQMEEWKKVANQLDKDHAK
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK  AGKLQSQLRTTVVAAAAFLDAFQKVADMAT  NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS  ALIDCLINPLQEQMEEWKKVANQLDKDHAK  EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRITTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRITTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRITTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRITTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
						THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM
1349	2699	A	10409	59	958	THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
						THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
						THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
						THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW

#### WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
  - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
  - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

# (19) World Intellectual Property Organization International Bureau



### 

(43) International Publication Date 9 August 2001 (09.08.2001)

#### **PCT**

# (10) International Publication Number WO 01/57188 A3

(51) International Patent Classification<sup>7</sup>: C07K 5/00, A61K 39/395, C12Q 1/68

C07H 21/04.

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875

3 February 2000 (03.02.2000) US 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US

09/496,914 (CIP)

Filed on

3 February 2000 (03.02.2000) 09/560.875 (CIP)

US

2000 (27.04.2000)

Filed on

27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R., Mintz, Levin, Coh. Ferris. Glovsky, and Popeo, P.C., One Financial Center. Boston, MA 02111 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, Fl, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, Cl, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

- -- with international search report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau
- (88) Date of publication of the international search report: 28 February 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



### INTERNATIONAL SEARCH REPORT

International application No. PCT/US01/05800

A. CLASSIFICATION OF SUBJECT MATTER							
IPC(7) :C07H 21/04; C07K 5/00; A61K 59/595; C12Q 1/68							
US CL :536/23.1; 530/300; +24/130.1; +35/6							
According to International Patent Classification (IPC) or to both national classification and IPC							
B. FIELDS SEARCHED	Live design washeld						
Minimum documentation searched (classification system followed	by classification symbols)						
U.S. : 586/28.1; 580/800; +2+/180.1; +85/6							
Documentation searched other than minimum documentation to the extent that such documents are included in the fields							
	the extent that such documents are included in the fields						
searched .							
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)							
Electronic data base consulted during the international search (	ame of data base and, where practicable, search terms used)						
STN, medline							
search terms: EST, expressed sequence tags	,						
C. DOCUMENTS CONSIDERED TO BE RELEVANT							
Category Citation of document, with indication, where ap	propriate, of the relevant passages Relevant to claim No.						
Category Citation of document, with indication, where ap	L. character and Land						
A VOLLRATH. D. et al. The Human Y	Chromosome: A 43-Interval 1-28						
Map Based on Naturally Occurring							
1992. Vol 258. pages 52-59, see whol							
1992. ( d. 200. pages 0 = 0 ), 100							
	· .						
	1						
	,						
, in the second of the second							
	<u>.</u>						
	_						
Further documents are listed in the continuation of Box							
Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand						
"A" document defining the general state of the art which is not considered to be of particular relevance	the principle or theory underlying the invention						
"E" earlies document published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step						
"L" document which may throw doubts on priority claim(s) or which is	when the document is taken alone						
cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined						
"O" document referring to an oral disclosure, use, exhibition or other means	with one or more other such documents, such combination being obvious to a person skilled in the art						
"P" document published prior to the international filling date but later	"A" document member of the same patent family						
than the priority date claimed  Date of the actual completion of the international search	Date of mailing of the international search report						
30 SEPTEMBER 2001	<b>16</b> NOV 2001						
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230  Authorized officer  JEHANNE SOUAYA  Telephone No. (703) 308-1235							
Commissioner of Patents and Trademarks Box PCT	IFHANNE SOULAYA MANA MANA						
Washington, D.C. 20231	The bare No. (70e) and there						
Facsimile No. (703) 305-3230	Telephone No. (703) 308-1235 /						